SYSTEMIC EFFECTS OF HUMAN FACTORS IN INFORMATION SECURITY

Timothy D. Kelley

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements for the degree
Doctor of Philosophy
in the
School of Informatics & Computing
and the
Cognitive Science Program,
Indiana University
November 2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

L. Jean Camp, Ph.D., Co-chair

Robert Goldstone, Ph.D., Co-chair

Alessandro Flammini, Ph.D.

Peter M. Todd, Ph.D.

28 July, 2014

This is dedicated to my parents and my partner in crime and minions: Kate, Harper, and Ada.

Acknowledgments

Due to limited time, space, and memory, it is impossible to thank everyone that helped or supported me. I should start out by thanking my committee. Imma let you finish, but I had one of the best committees of all time. One of the best committees of all time! They were very patient with me as I moved from bottom of the weeds to the top. Sometimes I can even communicate the big picture.

My family has also been a great source of inspiration. My partner's dedication to her own work and her own training provided me with additional perspectives when considering the types of questions to ask. My daughters have been a source of stress, but also stress release at the end of the day. There is nothing quite like coming home and being able to jump into a coloring book after a long day.

I would like to point out several important collaborators. Douglas Stebila convinced my advisor that eye tracking would be useful method for studying usable security. He also helped implement a lot of the lessons we learned in a fairly disastrous eye tracking pilot study and helped turned it into an insightful study. Ty Bross did the important programming for the mobile malware which is the basis for the final study in this dissertation. Without this work, it would have been a lot harder to believe the model presented there was realistic. Tom Busey was kind enough to let me tag along in his lab and use his eye tracking equipment. I would also like to thank both Andrew Adams and Kiyoshi Murata. They supported me for several months at Meiji University while we worked on an eye tracking paradigm to study potential cultural influences on digital awareness.

Communication with my peers has been invaluable. In particular, Grey Matters was a great resource for practicing my talks and discussing ideas. Nathaniel Husted, Roberto Hoyle, and Vaibhav Garg were instrumental in refining ideas for potential models and analysis. I would also like to mention Brandi Emerick, who taught me many of the basics of eye tracking which eventually made our study successful.

Finally, I would like to thank my revolving social circle. As a graduate student, one makes friends with many people at different stages and they leave and new people come into the picture. Being able to get together on Friday nights and let off steam and engage our imaginations, regardless of education and situation. I am glad it is something that we continue to this day.

Preface

It was just like a scene in an intrigue film
and I'm still not convinced that it wasn't for real.

This isn't intended for me, I don't think.
It's a missive from the edge of despair, I mean brink
of total desperation; the communication therein
says her hopes for survival are slim
and she's writing to the Front, though we've yet to meet,
with a confidential matter 'cause she's heard I'm discreet.
And the urgency of her request for my aid
is matched by the depth of the trust she displayed.

"Don't betray me like our oil minister did (staged a coup).
And I'm about to flee Nigeria soon
but I'll never make it out," she says, with twenty million
three hundred twenty thousand US dollars that are still in
her possession. She embezzled them, I guess.
Look, I don't really know her so uh... that's none of my business.

She's the LADY MARYAM ABACHA, deposed.
These days, can't even get her caps-lock key unfroze.
But yo, something about a widow in distress
(with 20 million dollars hidden in a metal chest)
softened up the Frontalot's heart, no doubt,
so I hit the reply button, tell her I can help her out.

She writes me back: DEAR FRONTALOT, UNITED STATES...
she acts so thankful. A bank full of money awaits!
And I hate delays so I'm quick to turn around
with my full name and the number to my checking account
and the scan of my license to drive an automobile
and my passport number, proving Frontalot's for real!
Then I'll meet the money in Stockholm. Ain't gonna walk home.

Think I'll retire to the south of Spain and sip gazpacho.
Not so quick, there's a little problem:
LADY A apparently had difficulty running all them
numbers I give her. But look, the fake ID's my only one,
and that's a real passport — I got it off usenet and checked,
I'm not dumb. I'm not some idiot
who's about to lose your money for you quicker than I'm getting it.

And of course my bank balance is negative; whose isn't?
That's why I need your 20% money laundering commission.
And I'm wishing I could talk about this further with you but I can't.
I just got an email from DR. UBUGU of Chad.
He's got a hundred and seventy-seven million in a bag.
I feel I got to help him 'cause his story is so sad.

MC Frontalot, Message 419.

Timothy D. Kelley

SYSTEMIC EFFECTS OF HUMAN FACTORS IN INFORMATION SECURITY

This dissertation couples the growing corpus of human subjects and behavioral research in information security with large-scale data and robust quantitative methods. Linking human subject experimentation with theoretical models enables the information security community to reason more effectively about the system-wide effects of user behavior. I examine how users interact with the digital environment, how those interactions affect decision-making, and how aggregate decision-making affects system-wide vulnerabilities. This interdisciplinary challenge requires a combination of techniques from cognitive neuroscience, social network analysis, human-subjects research, dynamical systems, network theory, and agent-based models.

In the first section, eye-tracking data demonstrates the relationships between expertise and online perceptual awareness of security cues. Expertise is shown to be only a small factor in attention to security cues, and task-type proves to be much larger indicator of attention, with tasks requiring the use of personal accounts driving attention to cues. This section uses Bayesian ANOVA to evaluate users' perceptual awareness of security cues as they complete common online tasks, as it relates to user sophistication and task type.

The second section uses a theoretical epidemiological model of malware spread to investigate factors that might mitigate the prevalence of malware in a coupled, two-population model. This both demonstrates that cost is the largest factor for affecting malware prevalence, outside of malware infection rates, and identifies appropriate strategies for system-wide botnet mitigation.

The final section utilizes an agent-based model of mobile application adoption combined with social network data and mobile marketplace policy. The result is an examination of the dynamic effects of user and market behavior on the spread of mobile malware and the second order effects, such as privacy loss, due to that spread. This model reveals that well-regulated markets are effective at limiting malware spread, but user behavior grows in importance as markets become less restricted.

Each study examines ways in which users interact with their technology, the aggregate effects of those behaviors, and identifies possible inflection points to change system-wide behaviors. This dissertation integrates empirical behavioral studies to develop a better understanding of digital behavior, thus enabling a more holistic approach to information security.

L. Jean Camp, Co-chair

Robert Goldstone, Co-chair

Alessandro Flammini

Peter M. Todd

Contents

1 Introduction

The growing corpus of human subjects and behavioral research in information security can be coupled with large-scale data and robust, quantitative methodologies from other disciplines to enable a more holistic approach to information security. In particular, linking human subject experimentation with theoretical models can help the information security community reason more effectively about the system-wide effects of user behavior. While traditional research and methods in information security will remain important for mitigating the technical risks found in devices, they fail to account for the dynamics of the social-technical systems in which those devices participate.

In this chapter I am going to briefly discuss the early evolution of information security and how the field opened up to include economics and psychology as important factors. Next, I will examine some of the interactions between cognition and design in users' behavior as it pertains to information security. Finally, I will briefly look at aspects of cognition, the effects of network structure, and social influence as they pertain to the diffusion of malware. I will conclude the chapter with a brief description of my proposed contributions as well as an overview of the rest of this document.

Early Evolution of Security Research

In the early days of computer security, the primary concern was the technical problem of access control to prevent the leakage of important, classified documents, to those that did not have the permissions necessary to view those documents. [210] This concern led to the development of numerous security models based on general systems theory where users and objects belonged to sets and relationships were developed to represent information flow in various state machines. [24]

As these models were analyzed, it became apparent that the security concerns of military systems were considerably different than those for commercial systems. [49, 59] There was also, fairly early on, the realization that computer usage, and networked computer usage would become a fairly normal occurrence and

the security models needed to take that into account. [57] And, while these models did consider economics, it was more about quantifying the extent of the problem, than a consideration of the problem in terms of economic theory. [59]

Market driven effects were found to be the primary source of reduced reliability and increased vulnerabilities. [7, 58, 111] These effects include that designing for security delays the deployment of software, allowing competitors to take advantage of network externalities due to users early adoption of the competitors' software. Early accounts of marketplace effects in security addressed issues of trust as an atomic element of a "trusted system" as opposed to a dynamic property developed through customer interaction with software manufacturers. [58, 60] they also examined the use of market forces on designing reliable software. [21]

The existing practice of secure software engineering, these early economic studies found, was not designed to account for evolving customer requirements. [33, 60] Furthermore, from a manufacturers perspective, producers will chose to remove, "as few defects as necessary and with as few early deaths as necessary," leading to software that is rushed to market full of vulnerabilities. [33, 60, 92] The long term presence of software bugs highlights the incentive misalignment between software development and secure software from a software manufacturer's perspective. [28, 40]

The study of the economic misalignments between software manufacturing and information security has led to numerous important findings. For example, the misalignment of incentives in software design creates lemon markets for secure products, where firms have no reason to purchase an expensive piece of security software because they cannot accurately assess its actual security. [6] Interestingly, the market for lemons also applies to criminal markets of stolen credentials, to user decisions about privacy, and, in a more direct connection with Denning's trust market, trust certified web-sites. [64, 101, 216]

Economic studies focus primarily on the behaviors between regulators and firms and have revealed some of the reasons that approximately 80% of computers remain unpatched. [184] For example, Choi, Fershtman, and Gandal demonstrate that there are ranges of behavior (e.g., if the cost of patching is too high) where firms will not apply patches, even if they have purchased software, and under some circumstances, will avoid purchasing products. [45] Similarly, work by Garg et al., point out the importance of economic factors in mitigating spam and spambots. [83] While useful, much of the work in economics of information security focuses on larger scaled actors, such as firms and governments, rather than the actors that compose these systems. And, when they do focus on the agents in the larger systems, they assume that the agents are

behaving in a rational manner. [8, 230]

Cognitive Security

The focus on large scale actors poses a problem for researchers in security and those that are implementing technical solutions to limit organizational risk, as the solutions that are put into place are often misaligned with the goals of the actors composing the organizations. [27, 40, 78] From an organizational perspective, this misalignment of incentives can create inadvertent insider threats, such as workers subverting security policy to get their jobs done. [27, 129, 205] It can also lead wide-spread vulnerabilities as users avoid secure behavior, such as software updates, to maintain a comfortable work/play environment at their residences. [205, 213, 220] In order to address these issues, a more comprehensive approach to understanding user behavior is needed.

In this section I will look at the interplay between design and users' cognitive abilities. I will begin by looking at some of the effects design can have on the cognitive processes of users. Then I will look at how mental models of security affect users' conception and responses to digital threats. Next, I will briefly discuss the nature of memory as a limiting factor in password use. Finally, I will look at phishing as an exploitation of trust, facilitated by the ease of misrepresenting–or users misunderstanding–aspects of digital identity. I will also discuss how expertise interacts with design heuristics as they pertain to phishing and limiting poor design decisions. [81]

Design and Cognition

Users' inability to utilize security features is due to at least two major aspects: poor design and the cognitive abilities of the users. [79, 167] Poor design can occur at the software level, leading to users being unable to find, let alone understand, the security features. [79, 180] I can also appear the organizational level, where policy requires users to choose between secure behaviors and completing their jobs. [100, 120, 197]

In both cases, poor software design and poor security policy design mitigates the benefits of the heuristic-based decision-making humans use when interacting with their environments. [81, 120, 134] In the case of poor software design, security features are distributed across many different menus or are so invasive users habituate to the warnings. [30, 198, 202] Poor organizational policy design can also habituate people to certain behaviors, such as offering up passwords to IT staff or creating weak passwords in order to comply

with a policy that requires users to change passwords too often.[1] [27, 100, 120] Poor design asks users to satisfice between performance–which users are responsible for–and security–which is mostly orthogonal to the task the user is performing. [188, 189]

Additionally, user interfaces can change rapidly, and further interact with additional changes in the operating system's user interface. These changes disrupt regularities, making it difficult for users to learn where to find normal features, let alone security features. [160,213] Humans rely on regularities–objects that co-occur in space and time–to learn in and from their environments. [207–209,237] Constantly updating the way that users interact with their digital environments exacerbates the statisficing that users are already performing by requiring them to spend additional mental resources re-learning how to do their jobs.

Technical sophistication and security aptitude can limit the negative effects of poor design, but this is context dependent. Experts have been shown to have improved working memory increased visual attention for domain specific tasks. [141, 232] This can mitigate aspects of a poorly designed system by allowing the expert to bring more resources to discover how the security measures function. [35] However, expertise should not be used to mitigate poor design, rather design should facilitate expertise. [36, 95]

While there are many technical solutions that limit the attack surface of a computer, users may be unable to utilize many of these features due to poor software design decisions. [107, 158, 226] Poor security policy decisions also place users in a position where they are willing to circumvent security in order to do their jobs. [27, 205] However, there are other cognitive aspects that affect users' decision making abilities in digital environments. In particular, users' mental models can affect how users approach security and the efforts they make in maintaining a secure environment.

Mental Models

In addtion to poor design, users often have the wrong sort of conceptions of malware behavior. Having an incongruous mental model of given security issue can negatively impact the course of action users take to resolve the issue. [17, 39] For example, users that conceive of malware as resource consuming and obvious fail to recognize that they are infected, leading to larger losses that are often translated across large distances

[1] As an example, a friend of mine working for a US bank in Japan told me that their security department implemented a new policy that required all email to be encrypted and that the encryption password had to be changed every month. My friend said that, within two months all of the encryption passwords were in one of two forms: <month><year> or <year><month>.

in time and space. [219] On the other hand, focusing on risk communication and solutions that match users' mental models increases understanding of risks, and limits inappropriate responses. [82, 219]

Additionally, Users' perception of authority in physical world does not transfer into the digital. Third-party trust seals–images that appear on websites to indicate that they have been verified as trustworthy by a trusted third party–have been shown to appear significantly more often on untrustworthy websites. [64] Users, however, rarely notice trust seals, and when they do, their trust of sites with seals is significantly higher. [116] The difficulty of translating physical risk perception into the digital is an important aspect of phishing attacks. [186]

Research in transfer of learning gives clues about how users use existing mental models and apply them to tasks that may seem dissimilar, but have many similar components. [56, 86] For example, communicating information security concepts in terms of con-men and home invasion to elder users was shown to increase their understanding of those concepts. [82] Similar techniques have been found to work to help users understand how firewalls and android permissions work. [128, 171] While communicating with mental models has been shown to increase understanding and motivation, there is little evidence that they fundamentally alter user behavior. [34, 82]

Moreover, transference of a physical mental model into the digital can have interesting repercussions. For example, a user with a folk model of hackers-as-burglar reported that, after they had been the victim of identity theft, installed an anti-virus scanner, effectively "changing the locks" and "putting bars on the windows", even though the information had been lost due to entering information into an insecure website.[2] Similarly, when looking at updates, users mental models affect how and when they install updates. In a study by Wash et. al., users that categorized their intentions for updating as "convenient" ended up with computers that, while enabled for auto-update, were left unpatched because the users were inconvenienced by the required restarts, or the effects the update downloads had on their Internet connections. [220]

Mental models affect how users perceive and react to digital risk decision making. When warnings and risk communication can fit a user's mental model, they have a better understanding of potential risks and are more motivated to act in a secure manner. [34, 82] However, when they are incongruous, users engage in much riskier behavior. [220] Mental models, cognition, and design also interact with how users choose and use passwords.

[2] Part of a conversation I had with Dr. Kami Vaniea.

Password Use

Passwords are still the predominant digital authentication method, but they suffer from numerous limitations. [102] In order to remain secure in the face of increasingly powerful password cracking techniques, passwords must become more complex. Moreover, users acquire numerous accounts over time. In order to reduce cognitive load, users choose weak passwords, reuse passwords, or write them down. [27, 84] There are tools for managing passwords, cued-recall systems improve users' ability to remember more complex passwords, and single-sign-on systems reduce the number of accounts users must maintain, but once again, poor design and misunderstanding lead many users to avoid them. [167]

Password cracking has continued to improve, from using clusters of CPUs to clusters of GPUs, requiring users to adopt more complex passwords. Researchers were able to crack roughly 57% of 173,686 unique passwords from compromised computers in a 24 hour time period. [200] A more recent attack using GPUs was able to search through SHA-2/512 passwords at a rate of approximately 356000passwords/sec. [63] Further modifications allow for grammar aware password cracking, which uses grammar rules and combinatoric dictionary searches to crack complex n-gram passphrases. [172] In fact, without a system that injects randomness into users' passwords, most passwords remain vulnerable. [222].

However, the passwords users select for their accounts are very limited. For example, the numeric sequences `1234` $\rightarrow$ `123456789` appear in the top 10 passwords on 4 different web pages that require accounts.[3]. [133] Similar regularities show up when users are asked to generate image-based passwords. [16] These similarities in image based systems have been shown to occur in a click-based cued-recall systems and theoretical systems that use eye-tracking for pattern recognition. [16, 75].

Additionally, the number of interactions that are required by users can limit their ability to engage in secure behavior. Users accumulate numerous online accounts over time, and, due to the number of accounts, they resort to reusing passwords or using physical mental aids–such as sticky notes–to help them remember their passwords. [109] Users employing a least-effort strategy combined with limited memory resources leads to a fixed-sized list of passwords and a growing list of accounts. [27, 84] This, in turn, leads to a high level of password reuse–approximately 3 accounts per password–which reduces the amount of work necessary to gain access to a compromised user's other accounts. This was demonstrated recently when

[3] `123456` was in the top position in 3 of the 4 websites, and the 3rd most used password on the 4th. Another favorite: `password`, was in the top 5 for 2 of the web pages, but did not appear in the top 10 for the other 2. The numeric sequences made up, on average, 64.7% of the passwords on the studied web pages. [133]

more than 40% of the passwords from a hashed password database were recovered using passwords found in the rockyou.com password database.[4] [133]

Despite these limitations, users seem to be reluctant to use technological solutions to help them manage their passwords. While it has been show that changes to the user interface can induce users to create more random and secure passwords, such changes are not widely implemented. [43] There are numerous ways to manage passwords, from browser-based password storage to software solutions dedicated to generating and storing passwords. Still users prefer to use their own memories, browser cookies, or paper notes. [84] Some of this can be attributed to users' perception of the problem:

> Anyone who wants to [compromise a password] can, you don't need to know them, and the shield of anonymity may make it less morally reprehensible to do so. [84]

If the attacker is going to break the password anyway, complex passwords are not needed, nor are devices to manage said passwords.

Users also have misgivings about using single-sign-on systems (SSO). Single-sign-on systems use an identity provider that vouches for the user to multiple service providers, effectively reducing the number of accounts a user must maintain. [23] However, due to perceived misalignments regarding user privacy and lack of open communication about what information is being transmitted to the service providers by the identity provider, users are avoiding SSOs. [23]

In addition to the limitations of complexity, number of accounts, and avoiding available technical solutions, passwords are also vulnerable to spoofing attacks. For example, a user can enter their password into a website that appears to be a legitimate website, but is, in fact, a malicious website masquerading as the intended website. The most common form of this type of attack is a social attack known as phishing, which manipulates users' trust relationships to entice them into revealing sensitive personal information such as passwords or financial information.

Phishing and Botnets

Phishing attacks mimic legitimate emails or websites to entice users to enter personal or financial information. [153] Phishing makes use of other forms of masquerading attacks, such as email address spoofing, and

[4] rockyou.com is an online gaming website. It was hacked in 2009 and all its 32603043 users' passwords were publicized. There were 14344386 unique passwords in the leaked password database.

users' misunderstanding of security features and lack of technical sophistication to appear as a legitimate entity, then use the appearance of authority and fear tactics to elicit responses from users. [229] In addition to collecting personal or financial information, phishing attacks are also used to launch drive-by downloads[5] to add the compromised device to a botnet.

Every aspect of the infected computer is economized once the device is added to a botnet. [146] The attacker can use the compromised computer to capture further personal and financial information by using keyloggers and man-in-the-middle attacks to intercept bank logins. [200] Phishing attacks also collect an infected computer's address books, which are either sold, or used to conduct context-sensitive attacks on the user's social network. [119, 165] The compromised device can also be used to host spoofed or fake websites, and to participate in spam campaigns. [147, 211]

Most users will not open an unsolicited email. [110] However, when attackers utilize contextual information, such as exploiting publicly available information collected from a target's social network, they increase the likelihood that an end-user will click through to the fraudulent website. [62] After they click through to the fraudulent website, users are generally unable to make sense of available warning signs. They rely on the look-and-feel because they do not identify aspects of URL spoofing, nor the intricacies of digital certificates. [181, 186, 191, 202, 234] Furthermore, these attacks exploit extant trust relationships, as well as potential fears of inaction, these attacks are effective against both naïve and sophisticated users, though sophisticated users are less susceptible. [106, 178, 186]

Phishing exploits users' social connections to facilitate the spread of botnets. The spread of malware and its connection to network structure has been the subject of much study. [196, 221] Early studies focused on the spread of malware through disk sharing, represented as a directed network. [114] This was followed quickly by an examination of the effects of a social connections in email would affect the spread of mail-based viruses. [157] Many of the obfuscation techniques botnets employ on compromised devices, such as day-night cyclical activity, have also been researched. [53, 54] However, to fully understand the social aspects of phishing, particularly phishing that takes place on online social media, it is useful to look at research that focuses on users' social behavior on networks.

[5] installing software either without the knowledge of the user, or by enticing a user to install a piece of software by falsely representing the nature of the software

Networks and Collective Behavior

Phishing and the growth of botnets has moved beyond the use of email. [94] In fact, the use of online social media has made it easier for botnet developers to entice users to visit malicious links, and, by using URL shortening, further hide the nature of those links. [42] Use of online social media as a vector for spam and phishing is between 2 and 4 orders of magnitude more effective than email spam and phishing campaigns, boasting a 0.13% click-through rate, compared to the approximate $0.0000081\% \rightarrow 0.00037\%$ found in studies of email spam and phishing. [204]

These studies focus on the interactions between the technical communication limitations–design–and the overall effectiveness of online social media spam and phishing. However, they fail to analyze the reasoning behind users' (both end users and malicious users) behavior. Examining how users participate in their online social networks will give us better insight into the factors that lead to such a disparity in effectiveness in spam campaigns.

Users generally have fairly regular digital behavior patterns. These include where they choose to visit on the Internet. While the overall structure of users' browsing behavior is heterogeneous in terms of what web pages users visit, there are local regularities. [76, 138] Users tend to choose the same websites to visit, and navigate primarily through bookmarks and favorite sites. [137] However, as users react to news events, they utilize search engine results to drive new link creation [137, 175]. Current cognitive research suggests that further contextualizing search results by creating social paths may help users aid one another as they navigate online spaces. [130]

User behavior is also context dependent, with users on Wikipedia navigating in a different manner than those using Facebook. [174] There are are also noted differences in behavior based on users' goals for communication. [173] This has been demonstrated in several domains relating to types of information search, national political discourse, and local organizational communication. [50, 51, 173] These contextual structures and behaviors, affect information spread and the formation of new connections, but they also create opportunities to identify abnormalities or to entrap users.

Attempting to manipulate traffic and behavior creates unique structures that can be analyzed by network analysis. This has been demonstrated in the manipulation of Twitter during political campaigns, the traffic patterns of botnet traffic, and the communication structures of criminals [97, 154, 173] However, malicious users exploit the regular browsing patterns and reliance on search engines to inject malware into search

results and advertising servers that host ads for many popular websites. [65, 123, 170]

While there are no current end-user technologies that exploit user regularities to issue warnings to users in the case of dangerous browsing behavior, there have been suggestions to develop very simple heuristics when dealing with potential threats due irregularities. [81, 156] Cognitive research into digital social search paths also shows promise. [130] Users can share information about potentially dangerous regions or sites, potentially limiting exposure to drive-by-download attacks.

This type of information sharing can utilize research in social learning. Recent research suggests that reliance on a peer group allows for innovation despite collective conservative behavior. [227] However, this requires methods to analyze the effects and popularity of given behaviors. [228] In terms of browsing behavior, a sample system has been proposed to allow users to share web histories anonymously with others, allowing users to leverage their social networks to limit potential reputation manipulation attacks. [101, 206]

Linking human subjects experimentation with theoretical models can help the information security community reason more effectively about the system-wide effects of user behavior. Cognitive neuroscience, social network analysis, human-subjects research, dynamical systems, network theory, and agent-based models are important tools to study how users interact with the digital environment, how those interactions affect decision-making, and how aggregate decision-making affects system-wide vulnerabilities.

Proposed Contributions

I use a combination of techniques from cognitive neuroscience, social network analysis, human-subjects research, dynamical systems, network theory, and agent-based models to examine how users interact with the digital environment, how those interactions affect decision-making, and how aggregate decision-making affects system-wide vulnerabilities. I provide 3 major contributions in my dissertation:

1. Using eye-tracking, I highlight the importance of task context in regards to users' attentional perception to security cues in a variety of online tasks. Task context, and its interplay with user sophistication, is an important first step in identifying the cognitive underpinnings of how users understand and react to potential risk in ditigal environments.

2. Using a new class of theoretical epidemiological model, I investigate factors that mitigate the prevalence of malware in a coupled, two-population model. My model demonstrates that cost is the largest

factor for affecting malware prevalence, outside of malware infection rates, and identifies appropriate strategies for system-wide botnet mitigation.

3. Using an agent-based model of mobile application adoption combined with social network data and mobile marketplace policy I examine the dynamic effects of user and market behavior on the spread of mobile malware and the second order effects, such as privacy loss, due to that spread. My model reveals that well-regulated markets are effective at limiting malware spread, but user behavior grows in importance as markets become less restricted.

Each study examines ways in which users interact with their technology, the aggregate effects of those behaviors, and identifies possible inflection points to change system-wide behaviors. My work in eye-tracking provides factors that affect attention to indicators of digital risk. The other two studies use previous work in digital behavior to construct and evaluate how different aspects of behavior collectively manifest.

My epidemiological work creates a framework for evaluating gross effects (e.g., cost, risk communication, risk-averseness) in an abstract population group. My agent based-model illustrates a possible manifestation of how those effects are expressed by examining the importance of user behavior as it interacts with smart-phone market place implementations.

Thesis Outline

My thesis attempts to point out important transdisciplinary work in information security by focusing on methodologies in cognitive neuroscience and complex systems, with a focus on behavioral research as a way to develop agent-based models for exploring the system-wide effects of actors in a given system. In Chapter 2, I will detail an eye-tracking study examining user skill and online tasks as it affects attention to browser security cues. Next, Chapter 3, will include a theoretical epidemiological model of malware spread that we use to examine important parameters for controlling malware spread. Penultimately, Chapter 4 will describe an agent-based model that explores the effects of marketplace implementation and user strategies for mitigating smartphone malware spread through social pressure. Finally, I will conclude by tying the various works together and suggesting future work and other methodologies for approaching the larger problem of ecology of information security in Chapter 5.

2 Competence and Online Perceptual Awareness

- Is there a competence-based observable, behavioral difference in how users perceive web browser security cues?
- More generally, what, if any, are the behavioral manifestations of expertise for online security?

In the case of eye-tracking, we find that expertise is not a significant factor in attention to online security cues, but task-type is. This suggests that both novice and expert users are aware of technical alerts to possible loss of security in web-browsing, but only employ them in certain contexts. Participants only paid attention to security cues when they used their own accounts. Our results tell us nothing about how experts and novices use the data to make decisions.

2.1 Introduction

Web browsers employ certain security indicators—such as the presence of the lock icon, the use of "https" in the location bar, or certificate information—to alert users to potential online threats, particularly related to the security of the web communications transmitted over the Secure Sockets Layer (SSL) / Transport Layer Security (TLS) protocol. Despite having been present in browsers for more than 15 years, many studies have demonstrated that security indicators are largely ineffective at communicating security information to users. This has been confirmed via self-reported usage of security indicators and eye-tracking data.

In other areas of security research, such as phishing, technical expertise has been shown to be a mitigating factor in user susceptibility to online attacks.

Recently, the nature of user authentication on the web has changed. While user authentication was originally site-centric—users had different usernames and passwords for each web site—the use of *single sign-on authentication* has allowed users to use their credentials from a single *identity provider* to log in to multiple sites, which are called *relying parties*. Single sign-on can provide several benefits to users: most

notably, they do not need to remember as many username/password combinations, and they do not need to register for new accounts at each site. On the other hand, there is a single point of failure (the identity provider), and users may have less control over their personal information.

Single sign-on systems are of growing importance within organizations and corporations. On the public Internet, several end-user single sign-on systems are currently available, both proprietary, such as the one provided by social networking site Facebook, and open, such as the distributed standard OpenID.

Web-based single sign-on typically involves authentication via redirection from the relying party to the identity provider; the user authenticates to the identity provider, and then the browser is redirected back to the relying party with authentication tokens which the relying party can use in a back-channel to obtain the user's profile information from the identity provider. Because of the redirection, information flow is much less clear than the traditional login process and may place a substantial cognitive burden on users.

In this work, we explore two related themes. First, we examine whether users with higher technical expertise make better use of security indicators in web browsers. Second, we examine to what extent users employing single sign-on make use of security indicators in web browsers and their degree of understanding of the flow of information in single sign-on. Our study employs eye-trackers to obtain data on actual user behavior, using both Facebook and OpenID as single sign-on identity providers.

Our goal is to provide answers to the following questions:

- Do users look for security indicators when using single sign-on in web browsers?
- Does the behavior of users with respect to security indicators differ between novices and those with computer or security expertise?
- To what extent do users understand the flow of information and risks involved in single sign-on? Do novices and experts have different understandings?

Approach

Our study involved 19 participants who completed a variety of online tasks involving both Facebook and OpenID for single sign-on and then filled out a survey. The surveys were used to compare reported behavior to observed behavior. Because eye-tracking is time-intensive, data-intensive, and often perceived as invasive, relatively small sample sizes are common.

While completing online tasks, participants' gazes were recorded using eye-tracking equipment. The online tasks included a variety of social networking tasks, such as rating an item on a movie website, sharing an item onto a social networking profile, and using a social networking account to login to other websites. Participants were asked to use their own Facebook account, but were provided with an alternative account upon request; for tasks involving OpenID, participants used a provided account.

The survey had three sections of questions: demographics, technological expertise, and single sign-on. Using answers from the technological experience section, we classified participants as either (a) novice, (b) computer experts, or (c) computer and security experts.

Results

After classifying users' expertise, we examined a variety of user behaviors and responses within the context of expertise. Here are some of the results of our analysis:

- Security experts have higher self-reported use of security indicators than non-security experts, and this is confirmed with eye-tracking data, both in terms of gaze duration and number of fixations at security indicators.
- Users with only computer expertise, not security expertise, have no more frequent self-reported or actual use of security indicators than novices.
- In general, users have a poor understanding of the flow of information during single sign-on. They do not understand the flow of credentials and profile information between the browser, the identity provider, and the relying party. They cannot correctly say whether relying parties learn the password for their account at the identity provider; computer experts are somewhat better than computer novices at this, though surprisingly we cannot say the same for security experts.
- Users do not always realize that they are using single sign-on, especially when doing so within the context of a single organization whose services are distributed across multiple internal web servers.
- Users *do* understand that, after logging in to a relying party via an identity provider, they need to logout of the identity provider when terminating their session at a public computer.

Outline

Section 2.2 reviews background on single sign-on, research on security indicators, and the role of expertise in security usability. In Section 2.3, we present the detailed methodology of our study. The analysis and discussion of our primary results is presented in Section 2.4. Additional discussion, including study limitations, appears in Section **??**, and Section **??** concludes. The study tasks and statistical analysis methodology appear in the appendices.

2.2 Background

Single sign-on

Single sign-on (SSO) protocols allow a user with an account at an *identity provider* to identify herself to a third-party service, called a *relying party*. Single sign-on can be used within a single organisation or across multiple organisations. Pashalidis and Mitchell [162] classified single sign-on systems into two classes: *pseudo-SSO systems*, in which a user authenticates to a single identity provider and the identity provider internally manages multiple credentials for that user to authenticate to relying parties; and *true SSO systems*, in which a user authenticates to the single identity provider but the identity provider only provides authentication assertions to relying parties. They also divide SSO architectures based on whether the identity provider component is *local* or a third party (called *proxy-based*). For example, the Kerberos protocol can be viewed as proxy-based true SSO system, whereas client-side public key certificates can be seen as a local true SSO system. We focus on proxy-based true SSO systems.

Only recently has single sign-on seen widespread implementation on the public Internet. OpenID [159] is a standard for federated authentication in which anyone can setup an identity provider and anyone can be a relying party, with no formal relationships required between relying parties and identity providers. Several commercial OpenID providers exist, and many webmail services also act as OpenID providers, but at present relatively few relying parties exist.

Closely related to single sign-on is the notion of delegated authorisation, such as in the OAuth protocol [1], where a user can delegate authority to a third party to access a particular resource on a server. For example, in August 2009, the popular microblogging site Twitter started requiring OAuth for all delegated authorisation.

In December 2008 the social networking site Facebook started offering a feature called "Facebook Connect" in which third party websites can allow users to login using their Facebook credentials rather than having to register for a separate account; this proprietary single sign-on service is built in part on the OAuth protocol.

For OpenID, Facebook Connect, and OAuth, single sign-on works via a sequence of redirects between webpages:

1. The user is on the website of a relying party, such as the movie review site Rotten Tomatoes.
2. The user clicks the "Login with Facebook" button on Rotten Tomatoes.
3. The user is redirected from Rotten Tomatoes to a Facebook login screen.
4. The user enters their Facebook username and password on the Facebook login screen and clicks "Submit".
5. Facebook verifies the credentials and asks the user to authorise the release of certain profile information.
6. The user consents to the release of profile information and then is redirected back to Rotten Tomatoes. The redirect includes cryptographic tokens that Rotten Tomatoes uses to subsequently request profile information from Facebook for that user.

Sun et al. [201] performed the first usability study of single sign-on protocols on the web. Participants using OpenID performed single sign-on related tasks using the existing browser interface and a proposed browser interface. They also surveyed attitudes towards single sign-on and comprehension of the risks and functionality of single sign-on.

Security indicators in web browsers

The Secure Sockets Layer (SSL) / Transport Layer Security (TLS) protocol provides encryption and authentication of communication on the Internet. The combination of SSL/TLS with web content delivered over the Hypertext Transport Protocol (HTTP) is jointly referred to as HTTPS. Authentication is performed using public key certificates a certificate authority (CA) who has verified that a given public key belongs to the legitimate owner of the given domain name and, in the case of extended validation certificates, that

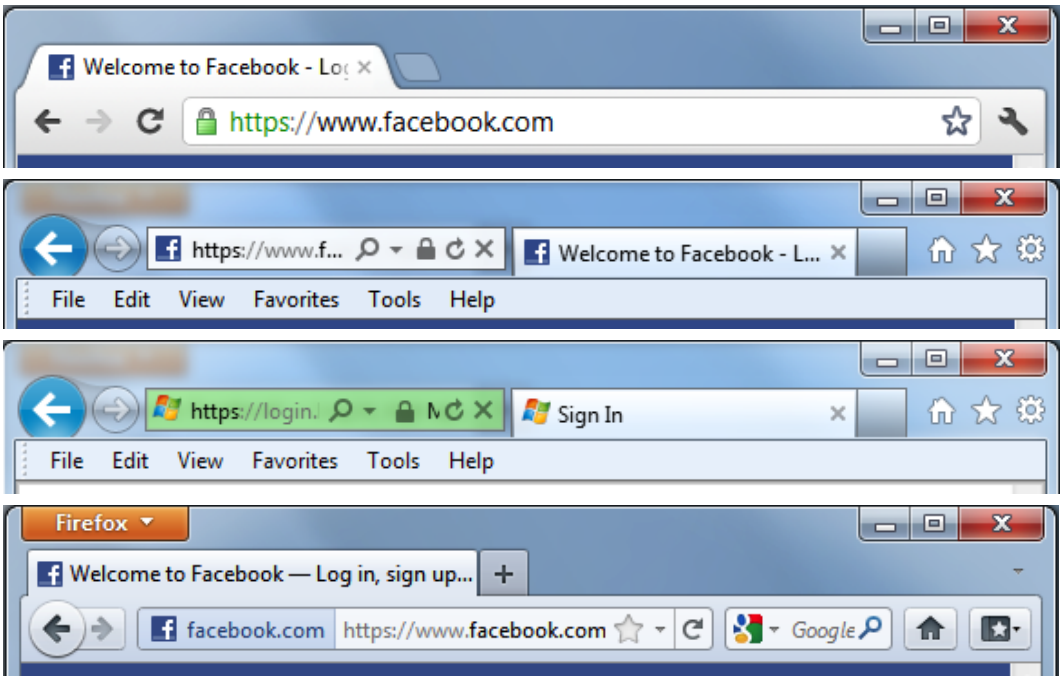


Figure 2.1: Web browser security indicators. Google Chrome 17.0, Microsoft Internet Explorer 9.0.5, Microsoft Internet Explorer 9.0.5 with an extended validation certificate, and Mozilla Firefox 10.0.2

the key belongs to that real-world entity or business. Multiple CAs exist, and today's popular web browsers typically trust upwards of 650 CAs [68].

Web browsers use several user interface elements, called *security indicators*, to denote the use of an HTTPS connection. Typically, these include the display of the protocol name (https) in the location bar, a lock icon, and additional coloring or elements for extended validation certificates. These security indicators are displayed within the browser *chrome*, the portions of the window controlled by the browser, as opposed to the *content* portion of the window that displays the HTML page. With several different major web browsers, different computing platforms, and frequent releases of new versions, the placement and semantics of security indicators in web browsers is inconsistent. Notably, Mozilla Firefox versions 4–13 did not display a lock icon to indicate the use of HTTPS, but it returned in version 14, after completion of our study. The security indicators for each browser used in our study are shown in Figure 2.1.

Research in the usability of security in web browsers traces its origins to the work of Friedman et al. [78] who conducted in-depth interviews to understand how users evaluate security of web sites. They collected types of evidence that users employed to decide if a website was secure: these included the above security indicators provided by the web browser chrome, as well as non-chrome indicators such as the type

of information requested, the type of the site, the quality of the site, and statements within the page about security.

Several subsequent works have investigated the extent to which users and websites employ these and other security indicators. Whalen and Inkpen [225] used eye-tracking equipment and interviews to analyse how users interact with security indicators: most looked at the lock icon, though few made use of its interactive capabilities to display certificate information; less than half looked for the the use of HTTPS in the location bar. Notably, no participants gazed at security indicators prior to being "primed" for security. Stebila [199] observed that popular websites do not consistently cause appropriate security indicators to be displayed; for example, of 125 popular websites, 19 had login forms delivered via an HTTP page but submitted via HTTPS, so no security indicators would be visible when entering passwords even though data submission would be secure; notably, this includes the world's second most popular website (according to Alexa Top Sites), Facebook, which is one of the single sign-on identity providers used in this study.

Certificate authorities, in conjunction with browser manufacturers introduced *extended validation (EV) certificates* in 2007; CAs would perform more extensive identity validation checks on parties (in exchange for more money), and browser manufacturers would introduce user interface elements, such as coloring the location bar green, to convey the purportedly greater trustworthiness of sites with EV certificates. Sobey et al. [191] analysed the relative effectiveness of the indicators of Mozilla Firefox 3 and their own modification; users did not generally notice the EV indicators in the standard Firefox 3, but their own modification was more successful. However, no major browser currently employs an interface similar to their modification.

Schechter et al. [182] observed that users will continue continue to login to websites when security indicators have been removed, and moreover even when security warnings are presented. Sunshine et al. [202] tested the effectiveness of various SSL security warnings; some designs were more effective than others, but in all cases a large proportion of users clicked through warnings.

Several works [166, 193, 194] have raised questions about the extent to which this insecure behavior can be explained by the artificial study environment. Complicating factors may include: participants using artificial credentials may feel less motivation to protect them; participants being "task focused"; and participants trusting that performing these operations in a study at a university means there is no risk.

Experts versus non-experts

Early research by Friedman et al. [78] observed that users from a high-technology neighborhood were better able to describe the security indicators associated with an encrypted channel compared to users from a less technical neighborhood. Sobey et al. [191] found that expert users were better able to identify extended validation certificate security indicators in web browsers. Sunshine et al. [202] in their research on the effectiveness of SSL warnings briefly consider whether technical expertise influences ability to identify warnings; they observed that experts made slightly better decisions than non-experts in some specific situations.

A real-life phishing attack performed by Jagatic et al. [106] on students at Indiana University found that students majoring in technical fields were roughly half as likely to fall for spear phishing emails as students in non-technical fields. Wright and Marett [233] confirmed that individuals with high self-reported computer self-efficacy or web experience, or participants who had high scores on a security awareness evaluation, were less susceptible to phishing attacks.

2.3 Design

In this study, we observed the behavior of study participants while performing certain social networking tasks; our observation equipment included eye-tracking devices to record where the participant's gaze was during the tasks. After the online tasks, participants completed a survey.

Participants were recruited via email and personal contacts; we aimed to recruit approximately 50% of participants as people we believed might end up classified as being security or computer experts and 50% as novices. Participants received a $15 gift card for participating, and could withdraw from the study at any time while still receiving the full value gift card, although no participants did. The study was conducted in a small computer lab at an off-campus university building. Descriptions of the study indicated to participants that we wanted to observe their of social media; we omitted any references to security in the study description or instructions.

The study was approved by the Human Ethics committee of the Queensland University of Technology and by the Institutional Review Board (IRB) of Indiana University.

Eye-tracker calibration

After signing the study consent sheet, participants were seated at a desktop PC running Microsoft Windows 7, with a widescreen 19" monitor (a trial version of this study run earlier noted that some eye-tracking systems perform poorly on small monitors). The PC was equipped with the Mirametrix S2 Eye Tracker, placed just below the monitor. This eye-tracking device has a data rate of 60 Hz with binocular tracking. The accuracy range of the device is 0.5 to 1 degree and the drift range is less than 0.3 degrees.

The device manufacturer's 9-point calibration routine was run. Accuracy varied by participant: participants without astigmatisms had average error of 40 to 50 pixels. Relying solely on manufacturer calibration had two drawbacks. First, the distance between some security indicators was less than 50 pixels, so it would be difficult to distinguish gazes at nearby indicators. Second, reported error for the device was averaged over the 9 calibration points, but subjects had differing inaccuracies: some users had small errors for points close to the centre of the screen but large errors for points near the edge of the screen, and vice versa.

See for example calibration errors for two different users in Figure 2.2.

As a result, we designed a secondary calibration phase in which we showed users additional calibration points which corresponded to points of interest for our study, for example the point at which the lock icon would appear when logging at a certain stage. We identified these points for each of the browsers in our study and prepared calibration videos. Participants were given a choice of web browser: Google Chrome 17.0, Microsoft Internet Explorer 9.0.5, or Mozilla Firefox 10.0.2 (the most recent versions of the browsers at the time of the study). We then showed users the secondary calibration video and directed users to gaze at the points in our calibration videos. This secondary calibration was done twice: before Facebook tasks and before OpenID tasks.

This secondary calibration phase allowed us to identify what the eye-tracker recorded for points of interest of our study, and compare those recorded points with points of gaze during the online tasks to determine whether participants gazed at our points of interest. Since the device's precision error was substantially lower than its accuracy error, this allowed us to obtain higher accuracy. We deemed a participant to have gazed at a security indicator while it was on screen if there was a fixation whose recorded distance to an indicator was within the average, plus two standard deviations, of the distances recorded for that point of interest during secondary calibration; two standard deviations ensured that all gazes recorded during secondary calibration would be accepted as valid, whereas one standard deviation would have missed some of

Figure 2.2: Eye-tracking calibration results for (a) a user with poor accuracy near the corners of the screen and (b) a user with poor accuracy near the centre of the screen

the calibration gazes.

Online tasks

After calibration, participants were instructed to begin the online tasks. Prior to the participant's arrival, we randomly decided whether the participant would be assigned to complete tasks involving Facebook first or tasks involving OpenID first. In this explanation, we will proceed for a participant who was assigned Facebook tasks first; the Facebook and OpenID descriptions below would be swapped for participants assigned to completed the OpenID tasks first.

Facebook

The participant was given the list of Facebook tasks as in the appendix. In summary: in task F1, the participant was asked to navigate to the movie rating site Rotten Tomatoes, login with their Facebook account, and rate a movie; in task F2, they were asked to post a story about a movie from Rotten Tomatoes to their Facebook profile and then log out "as if you were walking away from a public computer"; in task F3, they were asked to visit a blog on LiveJournal and post a comment on a story using their Facebook account, then log out; in task F4, they were asked to share something from Amazon on to their Facebook profile; finally in task F5 they were asked to log out, go back to Facebook, log in, and then log out one last time.

Each Facebook logins resulted in a pop-up window being displayed in the centre of the screen as shown in Figure 2.3. Participants were asked to use their own Facebook account if they had one, however they were instructed that they could request alternative credentials to use instead of their own. After logging in to Facebook, participants were asked to grant the relying party access to certain personal information. In the statistical analysis, we analyzed the login portion and the personal information grant portion of the tasks separately.

We did not alter the behavior or code of any of the websites used in this portion of the study. Although the front page of facebook.com is not delivered over HTTP, HTTPS is used to display the login page when Facebook single sign-on is accessed via a secondary website, so HTTPS security indicators were present during Facebook single sign-on login; no extended validation indicator was displayed as Facebook does not have an EV certificate. After logging in to Facebook, an additional page was displayed asking the user to share profile information with the relying party, security indicators may or may not be present on this screen depending on whether users have enabled Facebook's "Secure browsing" setting.

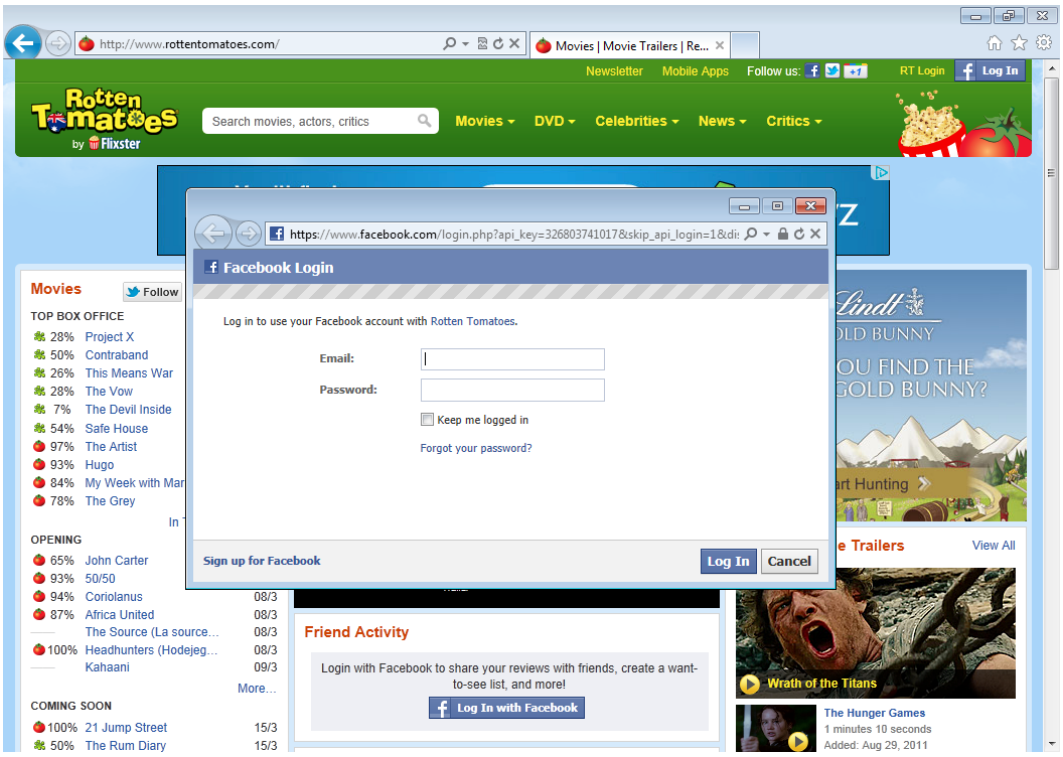


Figure 2.3: Login screen for Facebook single sign-on from Rotten Tomatoes in task F1 in Microsoft Internet Explorer 9.0.5

OpenID

The participant was next given the list of OpenID tasks as in the appendix. In contrast with the Facebook tasks, participants were not asked to use their own credentials. Instead, they were given credentials (an identity URL and password) for the OpenID provider we set up for this study.[1] In task O1, the participant was asked to visit a blog on LiveJournal and post a comment on a story using the provided OpenID account; in task O2 they were asked to visit a blog on BlogSpot and post a comment on a story, again using the provided OpenID account. Our OpenID provider operated entirely over HTTP, so no security indicators were ever present during interaction with our provider.

Survey

After completing the above online tasks, participants were given a 39-question survey to complete online. There were three components to the survey: (1) participant general demographic information, (2) information to assess the participant's computer and security expertise, and (3) information pertaining to their

[1] http://barnraiser.org/prairie

understanding of single sign-on and their behavior in the online tasks. Some of our questions were based on questions in existing survey instruments: section 2 included questions on technology expertise from Egelman [66] and Sotirakopoulos [194]; section 3 included questions on single sign-on comprehension from Sun et al. [201].

Some of the questions in our survey tried to identify which security indicators users use when signing in to websites. Where possible, we designed the sequence of questions in our survey to avoid priming participant responses: for example, in question 32 we asked the free-form question "How do you decide if it is safe to enter your username and password on a particular website?", but not until question 38, several screens later, did we explicitly list various security indicators and ask users to indicate which ones they used.

Upon completion of the online tasks and survey, participants were given a debriefing sheet with tips on using social networking sites, specifically Facebook, more securely. For participants that used their own Facebook account during the study, we offered to help them remove artifacts of the study from their account, including posts added to their wall/timeline and apps/websites linked to their account.

Classifying expertise

We used survey answers to classify participants on two dimensions: computer expertise and security expertise.

Computer expertise In the city and country in which we conducted our study, most people are indeed highly proficient at using computers. For example, no participants in our study answered below 3 on our 5-point Likert scale question (#9) where 1 was "I often ask others for help with the computer" and 5 was "Others often ask my for help with the computer". Thus, our rating of computer expertise was relative within this context. In particular, participants were assigned points for computer expertise as follows:

- 0.5: "Yes" to #8 "Do you use a computer daily for work?"
- 0.2–1.0: Answer to #9 "Rate yourself on this scale: 1—I often ask others for help with the computer . . . 5—Others often ask me for help with the computer"
- 1.0: "Yes" to #12 "Do you have a degree in an IT-related field?"
- 0.5 each: "Yes" to #13 "Have you ever. . . designed a website . . . created a database . . . written a computer program?"

The maximum possible score was 4.0. Participants with scores $\geqslant 2.5$ were classified as computer experts.

Security expertise We used answers from the following questions to assign points for security expertise as follows:

- 0.5 each: "Yes" to #13 "Have you ever... used SSH ... configured a firewall?"
- 1.0: "Yes" to #20 "Have you ever taken or taught a course on computer security?"
- 1.0: "Yes" to #21 "Have you attended a computer security conference in the past year?"
- 1.0: "Yes" to #22 "Is computer security one of your primary job responsibilities?"
- 0.5: "Yes" to #24 "Do you have an up-to-date virus scanner on your computer?"

The maximum possible score was 4.5. Participants with scores $\geqslant 2.5$ were classified as security experts.

Note that while the survey included several security-related free-form questions (#18 "If you know, please describe what a 'security certificate' is in the context of the Internet.", #19 "If you know, please describe what is meant by 'phishing'."), we explicitly did not use answers to these free-form questions in deciding security expertise. Instead, we used answers to the free-form questions to cross-check validity of the security expertise score above. Points for the free-form answers were as follows, up to 1 point for each question:

- #18 "If you know, please describe what a 'security certificate' is in the context of the Internet."
  - 0.5: Mentioned SSL or HTTPS.
  - 0.5: Mentioned use to secure communication or demonstrate trust of a website.
  - 0.5: Mentioned ownership of a public key.
- #19 "If you know, please describe what is meant by 'phishing'."
  - 0.5: Mentioned stealing user information.
  - 0.5: Mentioned fake email or fake website.

Eye-tracking data

During our analysis, we found that using eye-tracking data to answer the question "did the user look at this point?" is somewhat difficult. Users have a lot of eye movement during web browsing tasks, and may fixate near a point for just a fraction of a second; how long does the gaze need to be in order to "count" as having looked at that point? We will consider both the number of fixations over a security indicator and the duration of gazes at security indicators.

Statistical analysis

We analyzed the recorded eye-tracking data using Bayesian two-way analysis of variance and cross-validated our results using standard null-hypothesis testing. We examined mean gaze duration per fixation, mean number of fixations, and mean total gaze duration per task.

2.4 Results

Our study had 19 participants overall, but our eye-tracking equipment failed to record data for 1 of them.

During the online tasks, two participants requested to use alternative Facebook credentials rather than their own.

Note that based on survey question #33, few participants found the online tasks difficult, with no more than 3–4 participants (out of 19) rating any task "hard" or "very hard".

Participant demographics

Our participant pool consisted of 3 females and 16 males, with an average age of 26 and an age range of 18–39. Although our participants' gender and age distributions do not match that of the general population, several previous studies on Internet security suggest that gender and age do not affect participant security behavior [61, 194].

In terms of education, 1 participant had completed at most high school, 8 had studied some of or completed an undergraduate degree, and 10 had some postgraduate education. For those with some university education, 9 responded that they studied in a subject area related to information technology, 4 were in a subject area not related to IT, and 5 gave no answer. Only 5 of our 19 participants had English as a first language, but no participant appeared to have any trouble understanding instructions during the experiment.

All participants indicated that they had a Facebook account, and 17 had used their Facebook account at least once in the past month. Four participants with Facebook accounts reported having previously used their Facebook account to sign in to another website. No participant indicated that they had an OpenID account. However, given that many major webmail services are also OpenID providers, it is likely that many of the participants did indeed have an OpenID account but did not realize it.

In terms of web browser usage, during the study 9 participants chose to use Google Chrome, 2 chose to use Microsoft Internet Explorer, and 8 chose to use Mozilla Firefox. All but one user reported using one of the available browsers as their primary web browser.

Classifying expertise

Based on the methodology of Section 2.3, we classified participants on their expertise.

Per-participant results are displayed in Figure 2.4. Expertise values for each participant are aligned by Computer Expertise and Security Expertise.

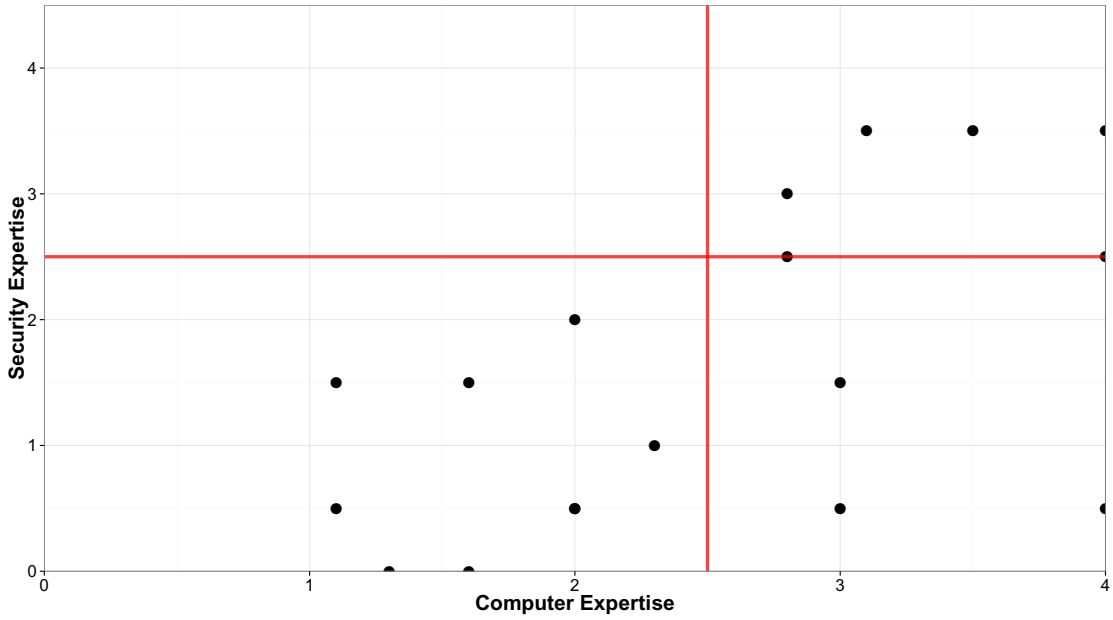


Figure 2.4: Subject computer and security expertise scores. Red lines indicate our cut-off values.

- *novices*: 9 participants had computer and security expertise scores < 2.5
- *computer experts*: 4 participants had computer expertise scores $\geqslant 2.5$
- *security and computer experts*: 6 participants had computer and security expertise scores $\geqslant 2.5$

No participants had high security expertise but low computer expertise. As a result, from here on we use "*security expert*" to mean "security and computer expert"; "*non-security experts*" includes both novices and computer experts who were not security experts.

To assess the validity of our security expertise questions, we included some free-form questions (#18, #19) in our survey, and scored participants on their answers to those questions as described in Section 2.3. We then compared the score on free-form answers to the classification based on the non-free-form answers. The mean score on free-form answers by the 6 security experts was 1.0, while the mean score by the 13 security non-experts was 0.423. The difference in means was statistically significant (Mann-Whitney $U = 62, n_1 = 6, n_2 = 13, p = 0.0398$).[2] We argue this provides evidence for the validity of our security expertise classification.

Use of security indicators

Whalen and Inkpen [225] were the first to analyze the use of security indicators use eye trackers. For a variety of security indicators, they compared self-reported use for each indicator with verification of use of that indicator via eye-tracking data. For example, in their study, 7 (out of 16) participants self-reported looking for https, and those 7 participants were verified as indeed having looked for https via eye-tracking; similarly, 12 participants self-reported looking for the lock or key icon, but only 11 actually did.

Eye-tracking evidence

As reported in Table 2.1, the majority of users in all expertise classifications did have a gaze point near the security indicators. We now explore in detail the extent to which task and expertise affect number and duration of gazes.

Number of fixations *Expertise effects.* When we examine the overall mean number of fixations between different expertise groups, we find that, while there appear to be differences between the groups in terms of number of fixations—with security experts having a higher mean number of fixations—those differences fall within our uncertainty measures and cannot be considered credibly different (Figure 2.5).

[2]The Mann-Whitney U test is a standard statistical tool that measures whether one set of independent observations tends to have larger values than another; in other words, whether the difference between two means is statistically significant. The test is *non-parametric*, meaning it makes no assumptions on the underlying distributions. Values of $p < 0.05$ indicate significance. Large values of p indicate that the data obtained does not demonstrate a statistically significant difference, though that alone does not prove the null hypothesis that the underlying behaviors are identical.

Table 2.1: Use of security indicators. Average total gaze duration (seconds) and average number of fixations on security indicators by task and classification for login and personal information grant dialog boxes.

	Security Experts (N=6)		**Computer Experts (N=3)**		**Novices (N=9)**	
Task	duration (secs)	# fixations	duration (secs)	# fixations	duration (secs)	# fixations
F1: Rotten Tomatoes/Facebook						
Login	1.788	2.667	0.679	1.000	0.794	1.556
Personal info	0.282	0.500	0.170	0.333	0.064	0.111
F3: LiveJournal/Facebook						
Login	0.110	0.167	0	0	0	0
Personal info	0.058	0.167	0.016	0.333	0	0
F4: Amazon/Facebook						
Login	0.036	0.167	0.186	0.333	0.274	0.556
Personal info	1.188	2.167	0.701	1.667	0.617	1.444
O1: LiveJournal/OpenID						
Login	0	0	0.455	1.000	0.142	0.444
Personal info	0.225	0.333	0.099	0.333	0.194	0.444
O2: BlogSpot/OpenID						
Login	0.126	0.333	0	0	0.049	0.111
Personal info	0.479	0.833	0	0	0.029	0.111

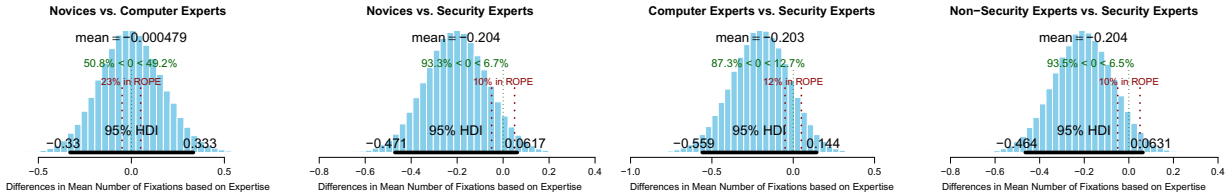


Figure 2.5: Differences in mean number of fixations based on expertise.

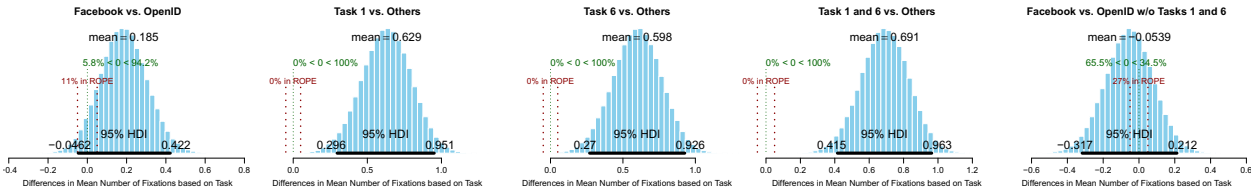


Figure 2.6: Differences in mean number of fixations based on task. In the charts, Task 1 refers to the login portion of the Facebook F1 task on Rotten Tomatoes, and Task 6 refers to personal information grant portion of the Facebook F4 task on Amazon.

Task effects. However, when we examine the data by task, a different story emerges. When we consider all of the Facebook tasks and compare them with the OpenID tasks we observe that while a mean difference of 0.0 is to the left of the distribution, it is fully within our 95% highest density interval, meaning that we cannot credibly conclude that the mean number of fixations differs in terms of task groups (Figure 2.6).

However, we note that Task F1 login and Task F4 personal information grant have noticeable and credible differences in terms of the mean number of fixations they receive. Furthermore, when we remove those tasks from consideration, any differences between Facebook tasks and OpenID tasks disappear (Figure 2.6).

Table 2.2: Results of ANOVA measuring Expertise and Task on mean number of fixations.

	Df	**Sum Sq**	**Mean Sq**	F value	Pr(> F)
A	2	5.10	2.548	1.647	0.196
B	9	70.96	7.885	5.095	5.85e-06
A:B	18	17.65	0.980	0.634	0.868
Residuals	140	216.67	1.548		

Cross validation These results correspond to our standard two factor ANOVA analysis looking at expertise (Factor A) and task (Factor B). Only Factor B was found to be statistically significant ($p = 5.85e-06$). Neither Factor A, nor the interaction were found to be statistically significant (Table 2.2). When we use Tukey's Honestly Significant Difference (Tukey HSD) to investigate the results, we find that only tasks F1 and Task F4 have any significant differences between the other tasks, supporting our initial Bayesian analysis.

These results demonstrate that users checked security indicators more during their initial login to Facebook, as well as when being asked to confirm sharing personal data on Amazon, but that overall number of gazes did not increase for the OpenID login tasks[3].

Gaze duration The number of fixations gives us a picture of what tasks subjects consider important, but it does not give us the full picture. We also want to know the length of consideration each subject gives to each fixation and the total time they spend gazing at security indicators. We analyzed mean gaze duration per fixation in two ways: First we look at a more fine grained model of expertise while only considering if the task is from Facebook or OpenID, then we look at a fine grained model of individual tasks, but treat subjects as either having security expertise (security experts) or not (non-security experts).

Expertise effects Looking at our fine grained model of expertise, we find that security experts, on average, gaze longer than novices[4]. However, our results suggest a bit of uncertainty: the mean difference of 0.0 falls outside our 95% HDI, but the 95% HDI includes part of the ROPE (Figure 2.7). A similar situation arises when we compare security experts with non-security experts. No mean difference lies outside our 95% HDI, but part of the ROPE is contained within the 95% HDI. We find no credible difference between security experts and computer experts, nor between novices and computer experts.

[3]Recall that OpenID login in our study occurred over HTTP, so no SSL security indicators were present. However, we still analyzed whether participants gazed at where those indicators would have been.

[4]Recall that "novices" excludes computer experts.

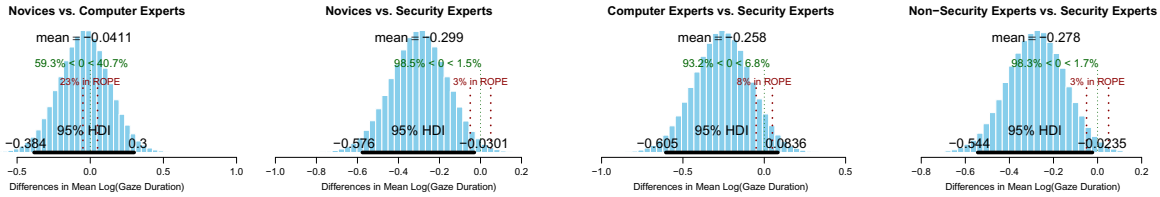


Figure 2.7: Differences in mean log(duration(secs)) based on expertise.

Table 2.3: Two-way ANOVA Analysis of Mean log(Gaze duration/fixation)

	Df	**Sum Sq**	**Mean Sq**	F value	Pr(> F)
A	2	2.049	1.0245	3.152	0.0472
B	1	0.352	0.3521	1.083	0.3005
A:B	2	0.429	0.2145	0.660	0.5191
Residuals	96	31.200	0.3250		

Cross validation When we cross-validate we find that the results of our Bayesian analysis are confirmed using a two factor ANOVA considering expertise (Factor A) and task type (Factor B). We find that expertise is an important factor (Table 2.3). However, when we use Tukey's HSD to exmaine the paired comparsions, we find that only the difference between security experts and novices is significant ($p = 0.061$), roughly corresponding to our Bayesian results.

When we compare security experts to those without security expertise (non-security experts)[5], the cross validation is stronger than our Bayesian results. The log transformation of the data gives us approximately normal data, allowing us to use a two-sample t-test to compare the groups. We used a Welch two sample t-test to compare the results due to differences in group variances. We found that differences between the groups were statistically significant ($\mu = -0.2858$, 95% CI $= -0.50284292, -0.06874345$, $t(99.869) = -2.6124$, $p = 0.01038$), confirming that security experts gaze longer than non-security experts.

Task effects. When we consider gaze durations per fixation based on expertise, and analyze the general task type differences we find there is no credible difference between Facebook and OpenID in terms of mean log(dwell time), meaning that subjects are gazing for roughly the same amount of time during Facebook tasks and OpenID tasks (Figure 2.8).

However, as we consider our fine grained task model, to see if any tasks receive more time per fixation based on expertise, we find that participants of all expertise levels gazed somewhat longer at security indicators during the login portion of task F1 and the personal information grant portion of task F4, but the

[5]Recall that "non-security experts" includes both novices and computer experts who are not security experts.

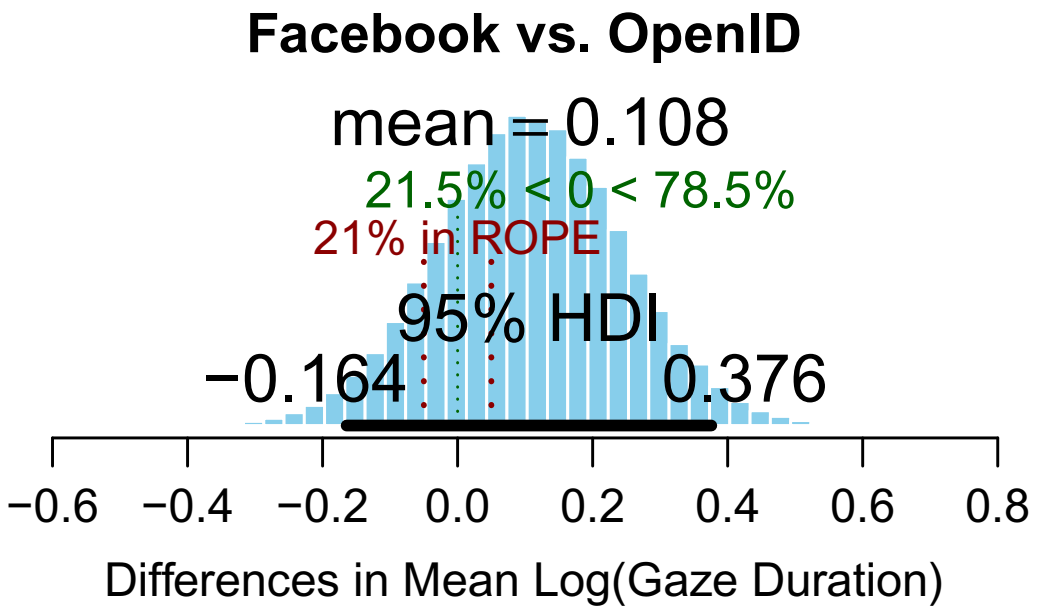


Figure 2.8: Differences in mean log(duration(secs)) based on task type.

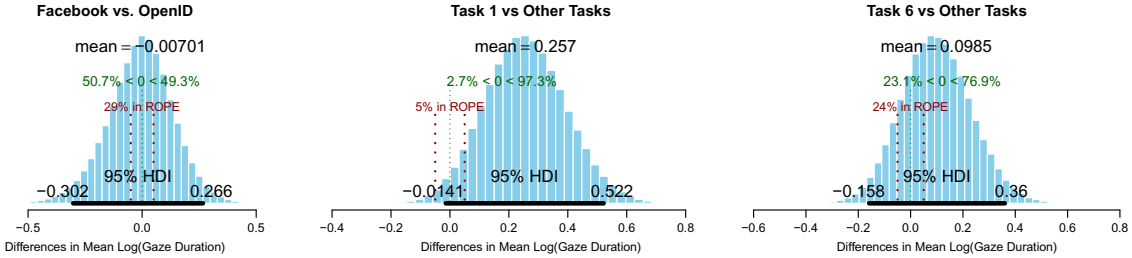


Figure 2.9: Differences in mean log(duration(secs)) based on tasks.

difference in average gaze duration during these tasks compared to other tasks fall within our uncertainty measures and cannot be considered credibly different (Figure 2.9).

Cross validation We we analysised the two factor ANOVA, we did find an statistically signficant effect for task on mean per fixation duration. However, in the post-hoc analysis, the only truely signficant differences were between task F3 and all other tasks. Looking at the data, this is due to an outlier effect, rather than any true artifact. No other tasks had any significant differences in mean per fixation durations.

Total duration Having both the number of fixations and mean per duration of fixation, we also wanted to examine differences between expertise (Factor A) and tasks (Factor B) on the total fixation duration per task, rather than per fixation. Due to time constraints and difficulty of specifying an accurate sampling

Table 2.4: Results of ANOVA measuring Expertise and Task on mean total duration of fixations per task.

	Df	**Sum Sq**	**Mean Sq**	**F value**	Pr(> F)
A	2	2.60	1.3018	2.193	0.116
B	9	24.49	2.7213	4.584	2.9e-05
A:B	18	7.29	0.4048	0.682	0.824
Residuals	130	77.18	0.5937		

distribution, we analyzed total duration using standard NHST techniques.

We find that task has a significant effect on total gaze duration, but there is still too much uncertainty about the effect of expertise to make a strong claim that it has an effect (Table 2.4). When we look at post-hoc Tukey HSD analysis we find that the only two tasks that are signficantly different from the mean are tasks F1 and F4, which makes sense, given that they receive more fixations total (Figure 2.6).

Discussion It is difficult to directly compare our rates of security indicator usage with those of Whalen and Inkpen [225]. Whalen and Inkpen simply reported the number of participants for which they "verified" that the user checked the indicator. In Table 2.1 we do report the proportion of participants that gazed over security indicators during login, but we see from the average duration of these gazes that the time spent looking at security indicators varies significantly. As a result, we cannot say whether our participants' observed use of security indicators matches or disagrees with that of Whalen and Inkpen. This does highlight the open issue of how to report and compare usage of security indicators from eye-tracking data.

Unlike Whalen and Inkpen, we do observe fixations without priming participants for security; they reported only observing no fixations before priming their participants for security. However, the nature of the fixations focuses primarily on tasks F1 (initial Facebook login) and F4 (share information with Amazon). The first represents the first time a subject logs in, while the second represents interacting with a commercial website. It appears that aside from these tasks, subjects give only cursory attention to security indicators, thus supporting Whalen and Inkpen's results that without security priming, subjects will not pay attention to security indicators.

On the other hand, our results seem to suggest that most users, regardless of security experience, are aware of security indicators and consult them in tasks they view as risky. This is an encouraging result, but, our experimental design does not allow us to determine the effects of the security indicators on decision making.

Table 2.5: Self-reported use of security indicators

Indicator	Security Experts (N=6)	Computer Experts (N=4)	Novices (N=9)
https	6	1	6
lock icon on the page	2	1	5
certificate	4	1	6
website privacy statements	2	3	4
type of website	6	2	6
professional-looking site	2	1	3
lock icon in browser	5	1	3

Other security indicators reported: "brand", "lack of ads", "firewall", "anti-virus safe browsing feature" ($\times 2$).

Self-reported use

Next, we compared self-reported use of indicators versus gaze duration. Recall that survey questions #32 and #38 asked subjects to report which security indicators they used to decide if it is safe to enter their username and password; in question #32, it was a free-form question, whereas in the later question #38 subjects were presented with a list of indicators and asked to check the ones that they looked for.

In the free-form question #32, only 4 subjects' responses mentioned one of the three accepted SSL security indicators (https, lock icon in browser, certificate); all 4 of these subjects were classified as security experts. (We emphasize that the classification in Section 2.3 of subjects as security experts did not depend on question #32 or question #38.)

In contrast, when presented with a list of security indicators in question #38, many more users reported using security indicators (Table 2.5). When prompted, participants self-report much higher use of security indicators. This may seem contradictory, but it demonstrates the limits of self-reporting and limitations of novices' abilities.

To analyze this data, we assigned each user a score between 0 and 3, with one point for each of "https", "lock icon in the browser", and "certificate", which are the only security indicators presented in the browser chrome. In particular, we omitted "lock icon in the page": as an element of the web page content, a lock icon on the page is not a trusted user interface element [199]. All but 3 participants reported looking for at least one of these three security indicators. The average score for security experts was 2.5, whereas the average score for security non-experts was approximately half that at 1.38, with the difference in these averages being statistically significant (Mann-Whitney $U = 62, n_1 = 6, n_2 = 13, p = 0.0408$). In contrast, the average score for computer experts was 1.8 compared to 1.67 for computer non-experts; the difference was not statistically significant (Mann-Whitney $U = 50.5, n_1 = 9, n_2 = 10, p = 0.6722$).

We then compared the self-reported use of security indicators with eye-tracking data. Of the 4 users that

self-reported use of security indicators in free-form question #32, 2 did gaze for security indicators and 2 did not. To analyze self-reported use of security indicators in prompted question #38, we used the security indicator score computed in the previous paragraph and compared it with security indicator gaze duration during Facebook login in task F1. The correlation was quite low and not statistically significant (Spearman rank correlation coefficient $\rho = 0.188$, $P = 0.4546$). [6]

Discussion The only statistically significant relationship we observed regarding self-reported use of security indicators was that, when prompted, self-reported use of security indicators by security experts was substantially higher than security novices. For all other relationships we considered, we observed no statistically significant differences. In particular, we observed no statistically significant difference between the security indicator gaze duration of security experts compared to security novices, or between participants who self-reported using security indicators and their gaze duration. Compared with previous results by Whalen and Inkpen [225], our participants had higher self-reported use of the https and certificate security indicators and somewhat lower self-reported use of the browser lock icon.

Understanding of single sign-on

To determine participants' understanding of single sign-on, we considered participants' successful completion of logout tasks and their responses to survey questions related to the flow of information in single sign-on. We asked participants questions about their previous use of single sign-on. Nine of 19 participants had heard of single sign-on, all of whom provided a reasonably correct definition (question #27); 4 were security experts, 2 were computer experts, and 3 were novices. Thirteen of 19 participants reported having "previously experienced using a single username and password to access different systems" (question #28). Of the 7 who responded "No" to that question, we have reason to believe as least 6 of them had in fact used single sign-on systems before, as they were or had been students at a university that the authors know uses single sign-on for a variety of services. This suggests that in today's web-based environment, users and system administrators do not have the same view of what constitutes "different systems".

Lack of awareness of using a single sign-on system may have implications on the way participants approach the given tasks. As we discuss in Section 2.5, many of our participants are opposed to single sign-

[6] Spearman's rank correlation coefficient is a non-parametric measure of dependence between two variables, in particular measuring how well the relationship can be described using a monotonic function. Values of ρ near ± 1 indicate a high degree of correlation, values of ρ near 0 do not.

Table 2.6: Participant logout actions on Rotten Tomatoes and Facebook in task F2

Classification	Logout of			
	RT&FB	RT only	FB only	none
Security experts (N=6)	3	0	1	2
Computer experts (N=4)	3	0	0	1
Novices (N=9)	3	3	1	2

on systems. It may be that opposition to single sign-on drives a more alert responses, if subjects are aware of the fact that they are using a single sign-on system. However, subjects that believe themselves to be using standard sign-on system may demonstrate more trust in the system, which is evident in less attention paid to security cues.

Logout

We directed the participants to log out several times: during Facebook tasks F2, F3, and F5, the participants were instructed to "Log out of the web browser as if you were walking away from a public computer. (Do not log out of Windows, however.)" Subsequently in task F5, they were asked to "Go back to the Facebook site and log in." and then "Log out of Facebook." All participants completed the last part of task F5—"Log out of Facebook". Participants' behavior at earlier tasks is more interesting; we focus on the first logout at task F2.

The participant is actually logged in to two websites during task F2: Rotten Tomatoes and Facebook. Rotten Tomatoes appears to make use of Facebook's single sign-on API for logout: users that log out via the link on Rotten Tomatoes are also logged out of Facebook. This is not a required feature of the Facebook single sign-on API, and users do not know a priori if dual logout will occur.

For task F2, we recorded whether users explicitly logged out of Rotten Tomatoes website and whether they explicitly logged of Facebook (Table 2.6). Overall, 14 of 19 users successfully logged out of either Rotten Tomatoes or Facebook, with 11 actually visiting Facebook to logout or check that they were logged out. There was no significant difference between the behavior of experts and novices.

We did not specify any logout task for OpenID, although our OpenID provider did have a logout function. Curiously, one participant did return—unprompted—to the URL for the OpenID provider to logout after task O2.

Table 2.7: Mental models of single sign-on

Classification		#34 Does Rotten Tomatoes know your Facebook password?		
	Correct drawing	**Yes** (wrong)	**No** (correct)	**Don't Know**
Security experts (N=6)	3	3	3	0
Computer experts (N=4)	0	0	3	1
Novices (N=9)	5	1	4	4

Discussion The participants seemed to demonstrate conservative logout behavior, in that when using a single sign-on service such as Facebook on a public computer, they logged out of both the relying party and the identity provider, and in particular all participants logged out of Facebook at the completion of the study.

Mental model

Several survey questions provide insight into participants' mental models of single sign-on, including the drawing exercise after question #33 and questions about password and profile information.

In terms of the drawing exercise, we used the same methodology as Sun et al. [201] for assessing the correctness of the mental model expressed in the drawing. As reported in Table 2.7, security experts did not do significantly better than security non-experts at answering this question.

Question #34 tested participants' understanding of the flow of information during single sign-on: it aksed if they believed that the relying parties, such as Rotten Tomatoes, learned their password for the identity provider Facebook. Table 2.7 reports the results: security experts did not do better than security non-experts, in fact they did worse.

Other

In January 2011 Facebook fully rolled out its "secure browsing" feature, which allows users to opt-in to having all Facebook pages delivered over HTTPS. Of the 17 participants who used their own Facebook account, we observed that only 3 had activated secure browsing.

After the initial eye tracking analysis we considered the problem with dealing with number and gaze durations due to random eye movement. While we did no correction for randomness, we were able to examine the eye gaze data in terms of Facebook use and OpenID use, which gave us a decent manipulation for random attention to security cues. With the Facebook tasks, dominated by tasks 1 and 6, participants

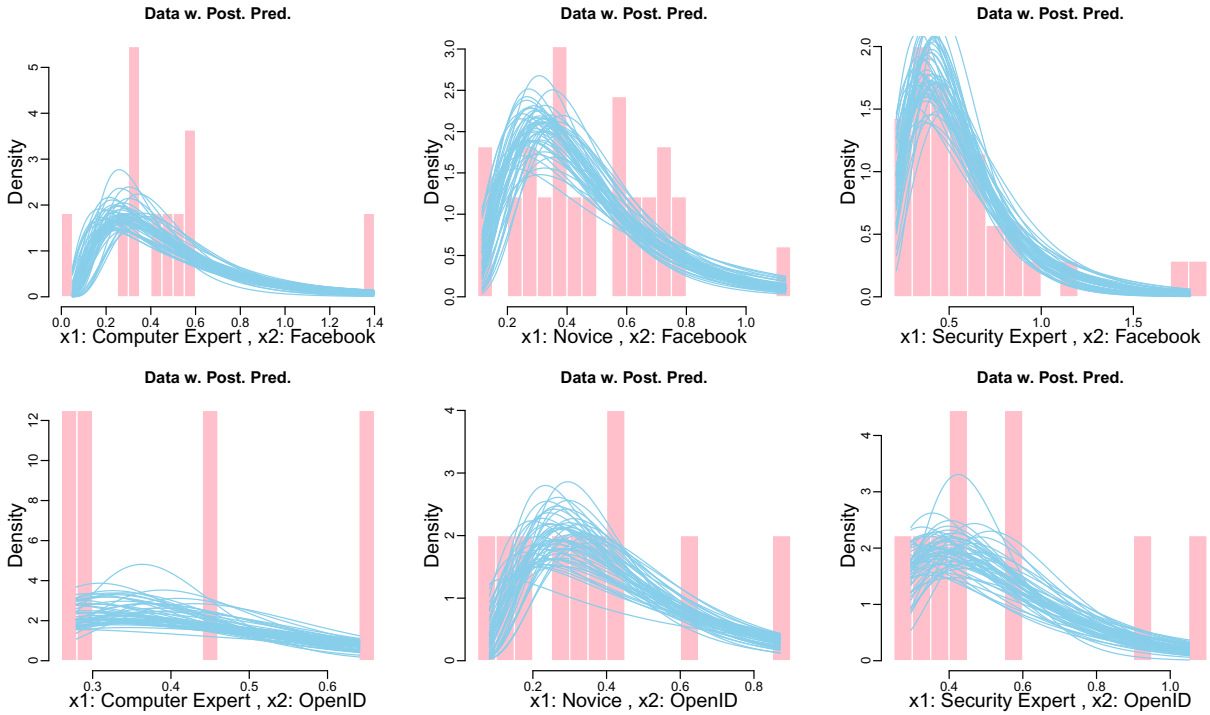


Figure 2.10: The posterior predictive results of the Bayesian model for fixation duration (blue lines) plotted over the recorded data (pink bars).

were generally using their own accounts, as opposed to the OpenID tasks, where no one used their own account.

Figure 2.10 demonstrates the difference in behavior between Facebook and OpenID tasks. While the full analysis showed suggestive differences between the two groups, seeing the resolution differences between the two tasks shows that while there are similarities in the means, the differences in variance (between Facebook and OpenID tasks) are indicative of the differences between random viewing and focused security-checking. Further analysis will directly examine fixations within security cue areas of interest (AOI) and further break down the fixations by task.

2.5 Discussion

In this section we discuss some additional observations. We examine participants preferences for single sign-on, the potential interactions between task and risk perception, and finally, our study limitations. We find that our participants are less inclined to use single sign-on, possibly due to privacy concerns. Similarly, we find that our participants pay attention to security cues both during initial log-on and when sharing

information with an explicitly commercial website.

Preference for single sign-on

When asked in question #37 if they would use their Facebook or OpenID account to login to third-party websites in the future, only 11% of participants responded "yes"; 42% chose "depends", and 47% chose "no". Contrast this with the results of Sun et al. [201], where they asked participants whether they would in the future prefer to use single sign-on in the form of OpenID (3%), in the form of Sun et al.'s identity-enhanced browser (9%), it "depends" on the type of site (36%, of which 30% preferred the ID-enhanced browser and 6% preferred OpenID), or not use single sign-on at all (29%). Our participants were substantially less inclined to use single sign-on than Sun et al.'s participants; since our participants only had the option of using traditional single sign-on as opposed to Sun et al.'s ID-enhanced browser, they did demonstrate a slightly more favorable response than Sun et al.'s participants did to OpenID.

Nature of task and risk

From our perspective in designing the study, the Facebook and OpenID tasks involved different levels of risk: in the Facebook tasks, most participants used their own accounts, whereas in the OpenID tasks all participants used manufactured accounts. However we noticed no significant difference in number of fixations or gaze duration between OpenID and Facebook tasks.

However, we did notice a difference between certain Facebook tasks. As noted in Section 2.4, participants of all expertise levels paid more attention to security indicators during their initial Facebook login, and when they granted Amazon–the only e-commerce site in our study–personal information. This suggests that users engage in a rational way regarding potential credential exposure, considering both personal accounts (Facebook tasks vs. OpenID tasks) and task types (login/information sharing vs. casual browsing) much the same way citizens chose whether or not to participate in the FTC's do-not-call list. [2, 32, 214]

The focus on security cues during login and commerce tasks also brings to mind consumer interactions with e-commerce. In regards to e-commerce, subjects are concerned with credibility and reducing uncertainty. [91] It is possible that participants are using the security cues as a measure of systemic reliability and credibility (*i.e.*, the security cues indicate a minimum of technical risk in the system) rather than indications of potential risks, particularly in the novice population group. [90, 98] However, these context dependent differences in security behavior require further study.

Study limitations and mitigations

It is well known that there are limitations to the ability of laboratory usability studies to reflect real-world environments [166, 193, 194]. We consciously made several study design choices aligned with recommendations previous work to try to reduce the impact of the study environment.

Setting. The setting of a study and the demeanor of the person running the study can have an effect on study participants. Individuals participating in a study—particular a security study—at a university can be of the frame of mind that "this is being run at a university, nothing can go wrong". Our ethics restrictions did not permit us to disassociate the study with the university, but we did take some measures to attempt to mitigate these factors. Our study took place in a university building a few blocks from the main campus, in an office tower in the city's central business district. The person running the study was a Master's student, casually dressed in shorts and a t-shirt.

In terms of the electronic "setting" of the study, we tried to match the participants' natural computing environment to some extent. Participants were given a choice of browser. All had previous experience using Facebook, so that single sign-on mechanism was not entirely foreign. One unavoidable unnatural characteristic was the use of eye-tracking equipment and the required calibration stage, though the device itself is relatively unobtrusive, and requires no further user attention once calibration is complete.

Demand characteristics refer to the "tendency for research subjects to guess the reason for a study, and then to attempt to confirm the experimenter's apparent hypothesis" [166]. All materials that our participants saw before and during the online tasks described the study as being interested in 'participants' use of social networking and social media', with no mention of security or privacy. Mentions of security only began in the survey, after completion of the online tasks.

Task focus is a risk in security usability studies: participants in studies are often highly motivated to complete the given tasks. Some previous studies [183] gave participants tasks to complete and then analyzed whether the participants completed these tasks even when security indicators or site authentication images were removed; participants who so completed the tasks were deemed to have not paid attention to indicators. Patrick [166] criticizes that approach due to task-focused participants being motivated to complete the tasks they have been given. As a result, we did not artificially remove any security indicators during our study, instead relying on eye-tracking data to assess participant attention to security indicators, both on tasks where security indicators were naturally present (single sign-on with Facebook which uses HTTPS), and

naturally absent (single sign-on with our OpenID provider which designed to use only HTTP). Moreover, our participants were promised that they would receive the full value of their compensation regardless of whether they completed the tasks or not. Nonetheless, in the informal discussions we had with participants upon completion of the survey, some participants reported task focus affecting their decisions.

Use of credentials. Schechter et al. [183] confirmed that study participants who use their own account credentials, rather than provided credentials, behave more securely. As a result, we asked participants to use their own credentials for Facebook tasks; we provided participants with alternative Facebook credentials if asked, which 2 participants did.

2.6 Conclusions

With ever more websites that users need accounts for, and with the growing popularity of social networking, the use of web-based single sign-on systems is likely to increase. With multiple parties involved—the user's browser, the identity provider, and many relying parties—users may have a hard time understanding what happens with their credentials and personal information, and what conditions should be satisfied for them to believe that a connection is secure or that it is safe to enter their username and password.

We examined users' use of security indicators in web-based single sign-on using Facebook and OpenID by employing eye-tracking equipment and surveyed users on their perception of information flow in single sign-on to determine if users with technical experts behave more securely than novices.

We found that users with security expertise did look at web browser security indicators more than those without security expertise; but computer expertise alone was not a predictor. Our participants—security experts and novices alike—in general had very poor understandings of the flow of information and trust in web-based single sign-on.

Important future work in this area includes the study of long-term trends. As users continue to use the Internet more and more and as their general computer proficiency advances, do they make better or worse use of security indicators? With the recent popularity of social networking, it seems plausible that web-based single sign-on will become far more prevalent in the coming years, and it will be interesting to see if and how users' understanding of web-based single sign-on improves as frequency of use increases.

3 User Behavior and Systemic Risk Reduction

- How can exploratory epidemiological models help us better quantify online risk?
- How can we use exploratory epidemiological models to help us better understand how online risk takers affect the security of risk averse populations?

We designed a model of two coupled populations, risk-takers and risk-averse, and explored how factors such as fear responses to malware, social alerts, and costs of maintaining risk-averse behavior affects the prevalence of malware. We found that a small-risk taking population poses a threat to a larger risk-averse population. Regardless of the effectiveness of malware, risk communication (social response) upon discovery of malware is the most effective method of reducing the prevalence of malware. Nearly as effective as risk communication, reducing the cost of maintaining a risk-averse behavior decreases the number of system-wide infections. Finally, a fear response to knowledge of a malware is effective initially, but not in the long term.

3.1 Introduction

Studying the spread of computer malware through the use of epidemiological models has been a useful tool in understanding the dynamics of individual outbreaks of malware, as well as giving some insight into possible mitigation policies. Kephart and White's early work focusing on system-wide prevalence examined effects of topology on virus spread as well the possibility of a social response to infection [113, 114]. Other work has focused on describing the dynamics of individual types of viruses, worms, or botnets.

In Kephart and White's examination of the social response, even a small social response was able to reduce significantly the total level of infection in the system. However, this result depends on a system where the recovered population could not become infected. For this simulation we wanted to examine the effects of social response when it led to recovery, but this recovery did not protect the user from reinfection.

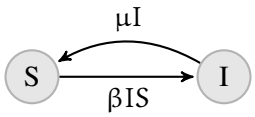


Figure 3.1: Permitted transitions in the SIS model.

We use the results from these individual models, as well as larger data on websites hosting phishing sites to model system-wide properties of malware spread. We use these system-wide properties to draw analogies from public health research regarding the spread of sexually transmitted infections (STIs) to examine organizational patching policies. From these results, we argue that thinking of security problems in terms of public health policy is a good addition to more traditional mental models of security.

3.2 Background

In the early work in adapting epidemiological models to computer viruses, the local nature of data transfer had to be taken into account. Computer viruses, in general, were spread very locally, and certain assumptions such as homogeneous population and the probability that an infected individual could infect any other individual in the susceptible population, did not hold [113]. In this environment Kephart and White (KW) adapted the SIS (Figure 3.1) model to a directed graph, to account for the non-homogeneous behavior of program sharing.

In their model, each computer is a vertex in a graph and an arc connects another computer in a program-exchange relationship. The arcs are associated with individual rates of infection and represent the set of vertices that can be infected by a given vertex, while each vertex is given an individual rate of recovery. Once a vertex has recovered, it is immediately capable of being reinfected. This, as the authors state, represents a very simple assumption that users will not become more vigilant after being infected. While this is a simple assumption, it seems to be a fairly good approximation for real world data [145].

Their deterministic calculations correspond to early results in prevalence driven epidemiological models [115], but failed to capture the social or organizational aspects of dealing with virus spreads. They modified their model to include a social response, or, as they call it, a kill switch model. That is, each computer, upon discovery and cleaning, alerts all other computers it is connected to to alert them of possible infection [114].

This extension (Figure 3.2) assumes that recovery corresponds to a temporary immunization from the

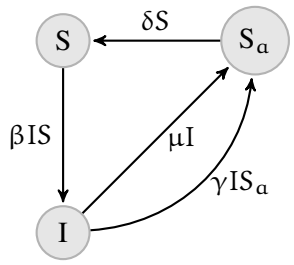


Figure 3.2: SIS model with Recovery and Social Response.

virus [114]. Based on these model extensions, KW show that central reporting and response to an incident is important to containing the incident. With central reporting and response, even if an organization is above the epidemic threshold, an incident can be limited in size and duration [114].

Other early work in modeling computer worms and viruses used fairly standard epidemiological models, in particular, variations of the SIR model. They overcame the limitations of the well mixed assumption by incorporating the scanning behavior found in many worms by expanding on Kephart and White's work on the effects of topology on virus spread. Knight, Elder, and Wang, analyzed networks in hierarchical and cluster topologies to study the effects of immunization from viruses in theoretical email networks [118]. Newman, Forrest, and Balthrop expanded Knight *et.al.*'s work by incorporating actual email network data and studies of network structure from the realm of statistical physics [157]. Both Knight *et.al.* and Newman *et.al.* demonstrated that targeted immunization could have a drastic effect on the spreading of viruses spread by emails.

Newman *et.al.* draws on Albert, Jeong, and Barabási's work on describing the network topology of the Internet [22, 236], as well as Pastor-Satorras and Vespignani's work on the effects of that topology on the dynamics of epidemic models [163, 164]. Albert, Jeong, and Barabási's work demonstrated the scale-free topology of the World Wide Web [22], but also the difficulty in generating models that reflected the true topology [236]. Pastor-Satorras and Vespignani, on the other hand, demonstrated that a scale-free topology could lead to the possibility of infinite duration, though low-level, prevalence of a given epidemic spread [163, 164].

In fact, we do see long-term endemic infections of older malware such as the Blaster Worm. [20] Bailey et.al. demonstrated that many older computers, which could no longer receive updates, continued to connect and attempt to infect other computers. This observable result, combined with Pastor-Satorras and Vespignani's work suggests that any piece of malware could have very long duration, and justifies our model

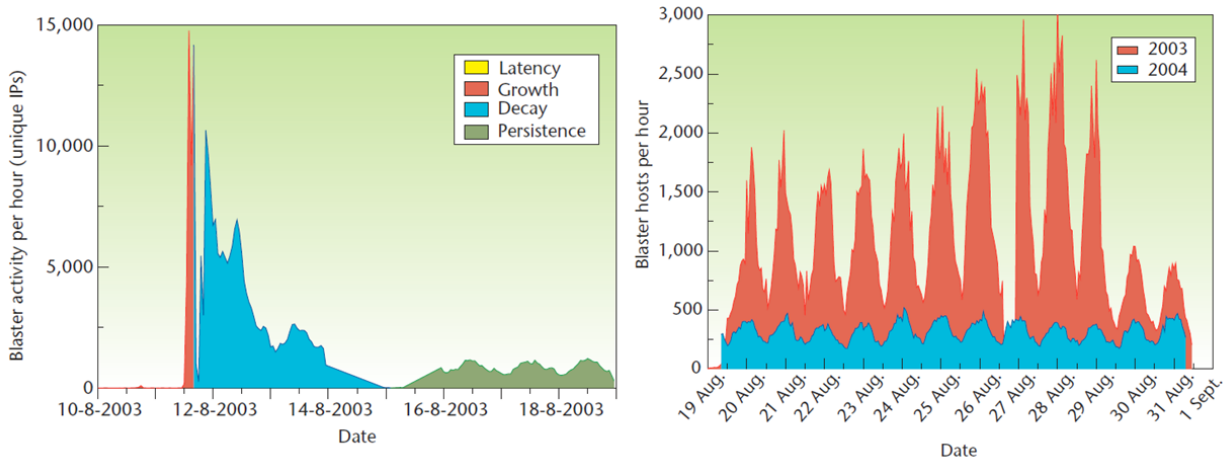


Figure 3.3: Epidemic and endemic dynamics of the Blaster Worm demonstrate that old malware can persist even after patches are available.

design decisions discussed in Section 3.3.

Zou *et.al.*'s work on modeling the Code-Red worm using the description and data provided by Moore *et.al.* modified the standard SIS model by incorporating a variation of Kephart and White's social response model, incorporating scanning rate, and allowing for infection rates to fluctuate in time [143,240]. Including the social response in their model allowed them to take into account human responses to the onset of an infection [240].

Zou, Gong, and Towsley, also included a model that allowed systems to become quarantined, removing them from the susceptible and infectious populations [241]. They demonstrated that removing computers from both populations for some amount time was an effective mitigating factor [241]. However, as Serazzi and Zanero point out in their later work on Sapphire, quarantines would be difficult to implement, as infected hosts cannot be trusted to quarantine themselves [185]. Zou and Towsley revisited their earlier work to demonstrate that the increased range of addresses in IPv6 would effectively reduce the total prevalence of routing worms such as Sapphire. This, they show, is due to scanning worms inability to access significant parts of the IPv6 address space in a reasonable amount of time [239]. Given that IPv6 space is currently and its allocation is optimized for reachability, this may not be true in practice.

Moore *et.al.*'s data collection and description of the explosive growth of the Sapphire worm required further modifications to earlier models [142]. While Code-Red generally followed standard models, Sapphire spread fast enough to become bandwidth limited, which in turn, limited its total ability to spread [142]. Serazzi and Zanero designed a model that encoded network resources. Utilizing incoming and outgoing

traffic rates into their model, they were able to capture the Sapphire's aggressive scanning. This scanning choked the Internet and greatly impeded Sapphire's rate of growth [185]. Serazzi and Zanero also point out the difficulty in implementing global security policies such as quarantines and hub immunizations.

Staniford, Paxson, and Weaver, contribute an excellent summary of many of the modeling attempts and call for a CDC for computer malware [196]. We agree the data collected via their suggested sensors and analysis would be useful for further mitigation of online pathogens. However, the focus of this chapter is more on the effects of risk takers on the total prevalence of contagion. Thus, we hope to show that a small group of users engaged in risky behavior creates a threat to the risk averse population.

To this end we look primarily at August and Tunca's work on allowing users with illegal copies of software to patch [18] and Choi, Fershtman, and Gandal's work on cost of patching [45]. While August and Tunca focus primarily on whether or not firms should allow users of illegal copies to patch, Choi, Fershtman, and Gandal look at the costs associated with different user's and their willingness to patch. We combine both the pirates in August and Tunca's work, with the non-patching populations of Choi, Fershtman, and Gandal, to show that limitations on user's ability to maintain a secure system is dangerous to the risk averse population.

Models of sexually transmitted diseases have become very complicated to deal the the multiple population interactions [**?**, 93]. However, most multiple population models do not couple the behavioral changes that occur do individual's perceptions of disease spread [168]. We build off of Perra *et al.*'s work to create a two population model with a social response that represents the ability of users to change behavior, and thus, their population group. This differentiates our model from more complicated models of STIs that use different characteristics of infection for individual population groups, but do not include behavioral responses to infection [19, 31, 176, 187].

3.3 Design

We first develop a simple model based on Kephart and White's initial social response model and Wang *et. al.*'s user vigilance model. We then use our model to examine the long term global prevalence of malware. Then we analyze the various parameters within this model to identify which ones are most effective at controlling systemic infection. We utilize anonymized data on websites used to support phishing attacks provided by Clayton and Moore, under a NDA, to demonstrate that the model can capture observed ma-

licious behavior. We also attempt to answer questions about feasible responses to malware diffusion that could result in reduction in botnet prevalence.

Model Creation

Kephart and White's social response model (KW) demonstrates the effectiveness of social responses to computer infection. We extend their model to allow the possibility of infection in the inoculated population. This extension includes aspects of Wang *et. al.*'s vigilance model [218]. Similar to their approach, we view user vigilance as prevalence based response to the infectious population, with vigilant users returning to the more susceptible population at a constant rate.

Wang *et. al.* assume that user vigilance declines after an initial peak due to responses to new infections [218]. Rather than using delay differential equations to model the decrease in vigilance, we make vigilance reliant on the infectious population, and separate vigilant and non-vigilant users into different compartments. Users return to the non-vigilant population at a constant rate, which we feel represents the effects of cost incurred due to maintenance of a secure system. Another option would be to have the return rate inversely proportional to the infected population: The larger the infectious population, the slower vigilant users return to the non-vigilant population.

Since we are modeling the diffusion of malware, the individual behaviors are of limited use. We are more interested in equillibrium states. Thus, the more common SIR type models previously used for models of malware infection are not useful for our purposes. This assumption better captures the long-term behavior seen in malware such as the Blaster worm [20] as well as the persistent insecurities found in web servers that allow them to be reinfected [145].

$$
\begin{aligned}
\frac{dS_r}{dt} &= -(\beta_r(I_r + I_a)S_r) - (\eta(I_r + I_a)S_a) + \delta S_a + \mu_r I_r \\
\frac{dI_r}{dt} &= \beta_r(I_r + I_a)S_r - \mu_r I_r - \mu_{a_1} I_r - \gamma_{a_1} I_r S_a \\
\frac{dS_a}{dt} &= -\beta_a(I_r + I_a)S_a - \delta S_a + \mu_{a_1} I_r + \mu p_2 I_a + \\
&\quad \gamma_{a_1} I_r S_a + \gamma_{a_2} I_a S_a + \eta(I_r + I_a)S_a \\
\frac{dI_a}{dt} &= \beta_a(I_r + I_a)S_a - \mu p_2 I_a - \gamma_{a_2} I_a S_a
\end{aligned}
\tag{3.1}
$$

The model we propose (Figure 3.4 and Equation 3.1) is a modified version of an SIS model with two interacting subpopulations. Our model does not assume immunity in the S_a population. This represents the

fact that no security system is 100% effective at stopping all vulnerabilities. We do not find, in the course of our analysis, that the rates of infection in the resistant population are so low that they may be ignored.

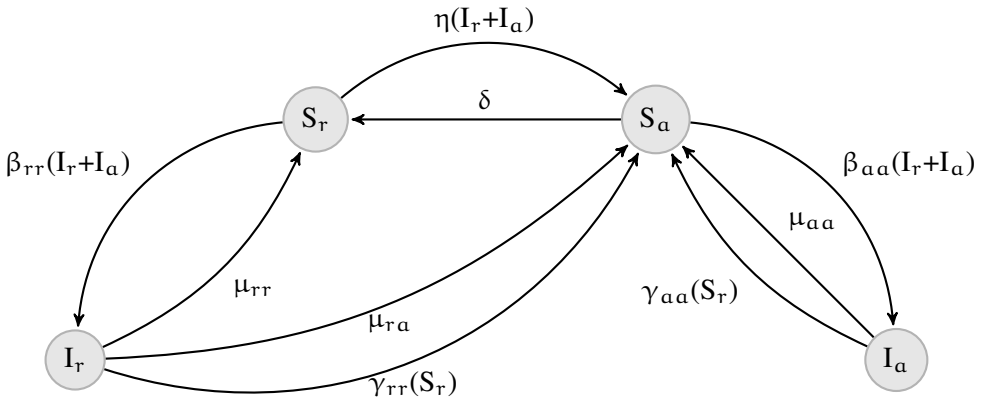


Figure 3.4: Two Population SIS model with Recovery and Social response.

Our model also assumes a well-mixed, homogeneous population. This is, in many ways, an unrealistic assumption, given the patterns of connection displayed by social networks and browsing behavior [138]. Moreover, it distracts from our metaphor of STIs, in that it assumes that all users are equally likely to interact with one another, rather than rely on contact patterns [74]. However, the dissemination of many online attacks is based on random scanning, which creates a scaled version of a well-mixed, homogeneous population [113]. Thus, this is a useful simplifying assumption, but can be expanded upon in future work.

Parameter Definitions

Table 3.1 briefly summarizes the various symbols we use in our model and analysis, which we describe here. S_r represents the susceptible population of non-vigilant users. These are systems that do not have a form of malware and can be infected. S_a represents susceptible systems within vigilant users. When an S_r or S_a system is infected, it transitions to the infected populations I_r or I_a, respectively.

η and δ govern the transitions between the two population groups. η represents the response of non-vigilant users to a given level of global infection. The higher η is, the faster non-vigilant users secure their systems. δ governs the response to the cost of maintaining a secure system. This is a constant rate, and the higher δ is, the less accepting users become of the cost, driving them to become insecure at a faster rate.

β_r, μ_r, β_a, and μ_{a_2} are the infection spread parameters for the non-vigilant and vigilant populations, respectively. β_r and β_a govern how fast an infection spreads, while μ_r and μ_{a_2} dictate how quickly a user

Table 3.1: Table Giving Definitions to included Symbols

Notation	Definition
S_r	Susceptible non-vigilant population
S_a	Susceptible vigilant population
I_r	Infected non-vigilant population
I_a	Infected vigilant population
η	Non-vigilant response to Infection
δ	Rate to return to non-vigilant population
β_r	Infection rate in non-vigilant population
β_a	Infection rate in vigilant population
μ_r	Non-vigilant recovery rate
μ_{a_1}	Non-vigilant to vigilant recovery rate
μ_{a_2}	Vigilant recovery rate
γ_{a_1}	Non-vigilant to vigilant social response rate
γ_{a_2}	Vigilant social response rate
R_∞	Equillibrium infected population
R_{∞_a}	Equillibrium infected vigilant population
R_{∞_r}	Equillibrium infected non-vigilant population

recovers. Recovery could be a simple as deleting an infected file, or as complex as reinstalling an OS. We assume that $\beta_r > \beta_a$ and $\mu_r < \mu_{a_2}$ to represent the fact that users that are maintaining a secure system will be less likely to become infected and more likely to recover.

γ_{a_1} and γ_{a_2} embed the response to social pressure to recover in the non-vigilant and vigilant populations, respectively. Users responding to these parameters, but not to μ_r or μ_{a_2} do not scan their systems for potential threats, but respond when an entity they know alerts them to a possible threat. For example, a user may respond to a Firefox reminder to update their browser or the exhortation of a friend. A specific instance of this was Google's effort to alert users to possible infections on their computers [122]. This is less than ideal for maintaining a secure system, as, with limited contact, infections can persist.

γ_{a_1} and γ_{a_2} are prevalence-driven in that they require some proportion of the population to be aware and able to fix the vulnerabilities that lead to infection. This means that without μ_{a_1} or μ_{a_2}, γ is ineffective at reducing spread of malware. γ provides a multiplier effect for the μ parameters, making individual maintenance of a system more effective. The use of γ in our model does not differ substantially from Kephart and White's use of γ, however, we do allow the risk-averse population to exert social pressure on the risk-taking population. Practically speaking, a user that pays attention to their system can catch an infection and alert their less aware neighbors. This allows a risk-taking population to take advantage of the effort from the risk-averse population.

μ_{a_1} defines the non-vigilant user's ability to clean or recover their system to a more secure state. This requires that non-vigilant users have access to the necessary patches and other up-to-date software to maintain a secure computer, at least until the cost of maintainence, δ, drives them back to the non-vigilant population.

Parameter Analysis

This model can be made equivalent to Kephart and White's "kill switch" model (Figure 3.2) by setting $\delta = 0.01$, $\beta = 0.5$, $\mu_{a_1} = 0.1$, and $\gamma_1 = 0.05$ and all other parameters to 0. We use both the Kephart and White (KW) model and a standard SIS model to compare our model under different parameter conditions. This allows us to evaluate which parameters may be realistic and useful.

Parameter Analysis in Vigilant Population Only

We first analyze the various effects of adjusting the parameters on the vigilant population to identify the most important parameters in controlling infection in that population. From there we move to analyzing the whole system, individually adjusting certain parameters to identify the key parameters in the system as a whole. For these simulations we vary one parameter and keep others constant. For each of the parameters we hold constant: β_a, μ_{a_2}, and γ_{a_2}, we set them to $0.5, 0.1, .01$ respectively.

These parameters are taken directly from KW and varied in later simulations. This sets a fixed social response at 1/10 the level of the cleaning response. This allows us to maintain consistency with our system-wide analysis below. We then vary the parameters of interest for each simulation from 0 to 1 by .01. Because we are only working with S_a, we initialize the populations to: $S_r = 0, S_a = 0.99, I = 0, I_a = 0.01$. Without an infected population I_r or I_a, no further infections are possible in this model.

Parameter Analysis with Both Populations

The system parameter analysis keeps the infection rate and cleaning rate in the non-security aware population at the same level as the standard SIS model used by KW ($\beta = 0.5$ and $\mu_{r_1} = 0.1$). This reduces the number of variables we must examine, and provides us with a reasonable worst case scenario of eighty percent of non-vigilant computers infected. However, we adjust the security aware population to reflect a greater vigilance.

We set the infection rate of S_a to half of the non-vigilant population's rate. Similarly, the cleaning rate of S_a is twice that found in the non-vigilant populations. In the vigilant population, there is a social response,

but this is 1/10th the cleaning rate. This leads to the following parameter values: ($\beta_a = 0.25$, $\mu_{a_2} = 0.2$, and $\gamma_{a_2} = 0.02$). We normalize the initial populations to $S_r = 0.99, S_a = 0, I = 0.01, I_a = 0$.

Moreover, these parameter values are a reasonable estimation of actual global prevalence. Our intial parameter values in isolated populations lead to roughly 80% of the population falling into the non-vigilant population, and roughly 80% of that population infected. Within the vigilant population, the initial parameter values lead to roughly 13% of that population infected. With no interactions between the populations, this leads to a global prevalence of roughly 77%. These results correspond to estimates of global prevalence.

In their report to the House of Lords in 2007, the Science and Technology Committe reported on results from an earlier study that showed that roughly 80% computers lacked neccesary security measures, and roughly 72% of sampled systems had some type of malware [184]. However, the committee noted that this study only sampled 354 computers, so it probably was not an accurate portrayal of the actual prevalence of malware. The Anti-Phishing Working Group's (APWG) mean observed infection rate for 2010 and the 1st half of 2011, which is approximately 48% [10–12]. Thus, our initial parameter values align very closely with the earlier study, and represent approximately a 60% increase over the APWG's results.

Uncertainty and Sensitivity Analysis

The first two sets of analysis represent a very crude sensitivity analysis given the number of parameters. We analyze each parameter in light of a fixed system. This reduces the problem from a nine dimensional problem to a one dimensional one. However, it is not informative in terms of how the parameters interact with one another. To address this, we performed two types of sensitivity analysis in two situations.

We applied used Latin Hypercube Sampling (LHS) on the set of all parameters to measure both epistemic uncertainty and to perform a sensitivty analysis of the parameters given the measured value of total infectious computers [29]. LHS first samples from prior distributions of parameter values and generates sampled output for the number of samples. In our case, we used 1000 samples. This gives us our uncertainty analysis as we examine the variability of outputs as we vary the parameter values. From there LHS use a rank-transform correlation coefficient to measure the sensitivity of each parameter as it pertains to the measured output [29].

We used uniform priors for all parameter values, given our own uncertainty of acceptable distributions. Our initial test was performed over all parameters and used to identify the key bifurcation parameters [135]. These bifurcation parameters are key to differentiating the major equillibrium behavior in the system; mainly, whether a contagion is maintained or dies out. After we identified the key bifurcation pa-

rameters, we set them to ensure continued prevalence and performed the analysis again. This allowed us to do uncertainty and sensitivity analysis of the secondary parameters associated with reducing prevalence.

Model Fitting

To fit our model to data we used MATLAB's `lsqcurvefit` function to fit the cumulative sum of the infected risk takers and risk adverse populations to the cumulative sum of the observed data. All parameters, as well as the initial population values, had a lower bound of 0 and upper bound of 1. The population values were normalized to 1 before computation. We took our total population to be 150,000 based on the cumulative sum of the total number of attacks. We ran each fit 50 times to try to avoid local minima.

We fit our model to the top ten companies targeted by phishing scams. This data contains observed websites spoofing legitimate businesses such as banks and other online commerce. This does not represent the social aspect of the attack, but rather the infrastructure used to support such attacks. We also fit our model to the total number of attacks for those companies.

To identify the top ten targeted entities, we cleaned the data to attempt to get unique identifiers. We found that, for the most part, our cleaning manage to capture most of the necessary data, but it can be refined for a more accurate picture. However, that being said, the top ten targeted entities account for 130559 out of 157355 observed attacks, roughly 83% of the total observed attacks.

3.4 Results

We examined the effects of parameter variation in different phases. We first wanted to see if there was a way to reduce total infection prevalence by adjusting only the parameters associated with the vigilant population. After considering only the vigilant population, we investigated the effects of making the non-vigilant population respondent to the vigilant population.

We conducted this investigation by adjusting the μ_{a_1} and γ_{a_1} parameters to investigate the effect of a user's ability to recover to more secure behavior. We then adjusted the parameters that determined the speed of transition to and from the vigilant population (η and δ) in uninfected users. When the infection is less potent in vigilant users, we can reduce the total infected population by having more users become vigilant and having vigilant users stay vigilant longer. For these simulations we do not adjust the infection or clean rates, but keep the non-vigilant and vigilant population parameters at their fixed rates discussed above.

Effects of Adjusting Parameters in Vigilant Population Only

In KW's model of social response, they add a prevalence driven recovery effect on top of the standard, constant rate recovery. In order to investigate the effects of this recovery in the population we varied the social response in the vigilant population only, to see if it would lead to significant reduction in the equilibrium of total infected. Since we split the model into two subpopulations, we could also examine source of the infections.

Simulation 1

In Kephart and White's examination of the social response, even a small social response was able to reduce the total level of infection in the system. However, this relied on a system where the recovered population could not become infected. For this simulation we wanted to examine the effects of social response when it led to recovery, but this recovery did not protect the user from reinfection. We set the infection characteristics in the vigilant population to correspond to KW's model ($\beta_a = 0.5$ and $\mu_2 = 0.1$) and varied γ_{a_2}.

Looking at the results (Figure 3.5) we find that only when the combination of social response and cleaning rate is greater than the infection rate does the infection die off. That is, $\frac{\mu_{a_2}}{\beta_a - \gamma_{a_2}} = 1$ is the bifurcation point in this single population. When $\frac{\mu_{a_2}}{\beta_a - \gamma_{a_2}} < 1$ then $R_\infty = 1 - \frac{\mu_{a_2}}{\beta_a - \gamma_{a_2}}$, and when $\frac{\mu_{a_2}}{\beta_a - \gamma_{a_2}} > 1$, the infection disappears.

These results mean that the total social response and the cleaning rate must effect the network at the same rate as the malware to be effective at eliminating its spread. For a single population then, social response is a useful measure in reducing the total infection rate, but is unlikely to be able to reduce the infection from a pandemic unless either the cleaning rate or the social response is unreasonably high. For example, the dynamic characteristics of Conficker.C described by Antonakakis *et.al.*, could be considered a pandemic infection [13]

As we will see, when we look at Simulation 9, the social response that allows non-vigilant to become vigilant, γ_{a_1} is an key ingredient to reducing the total R_∞. In our model a shift for the vigilant to the non-vigilant population is associated with an increase in the rate of social patching. An increase in the expense of recovery decreases the vigilant population and thus decreases the presumptive efficacy of social recovery. Moreover, since γ_{a_1} is prevalence driven and δ is constant, any amount of shifting is able to produce some amount of prevalent vigilant population.

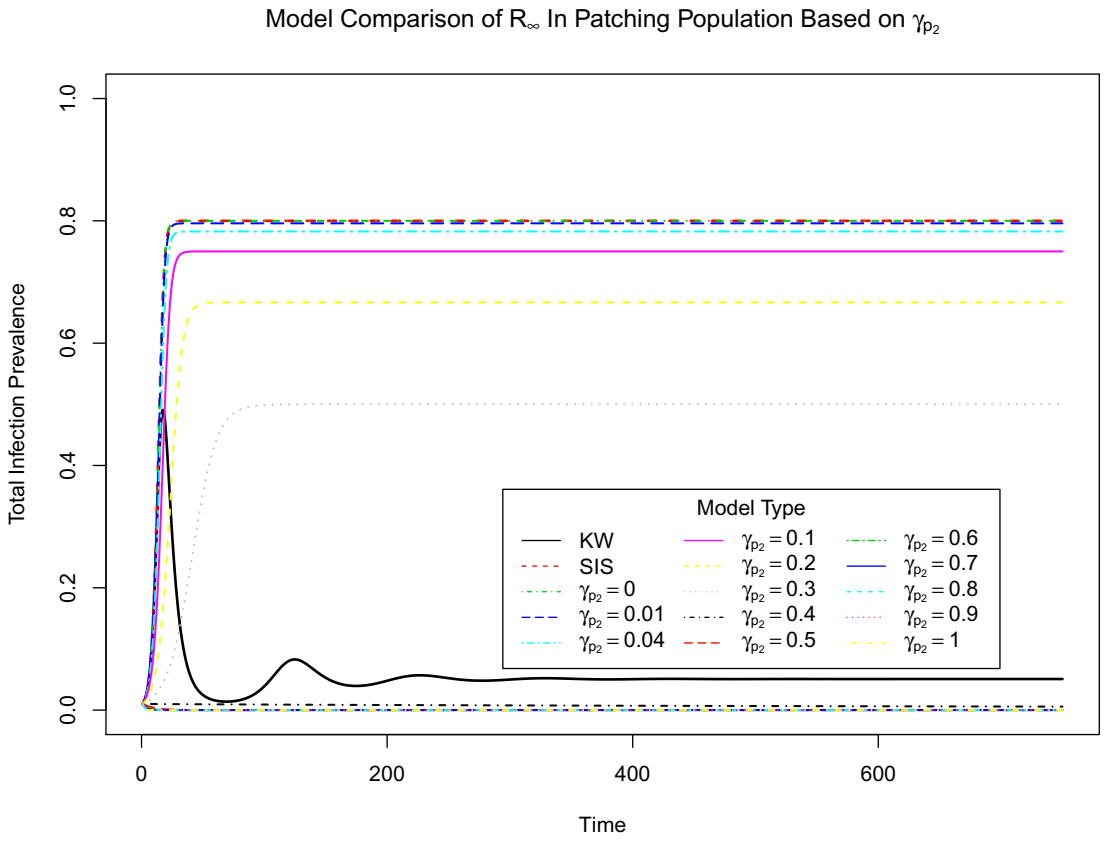


Figure 3.5: Comparison between SIS, KW, and Our Model, γ_{a_2} variable. Only when γ_{a_2} reaches high rates relative to β_a does R_{∞_a} fall.

Simulation 2

In Simulation 2 we varied the cleaning rate in the vigilant population. Without any cleaning rate, even in the presence of a social response, $R_\infty = 1$. In fact, when $\mu_a = 0$, the bifurcation point becomes $\frac{\beta_a}{\gamma_{a_2}} = 1$, and when $\frac{\beta_a}{\gamma_{a_2}} > 1$, $R_\infty = 1$. When $\frac{\beta_a}{\gamma_{a_2}} < 1$, $R_\infty = 0$. This means that, within a given population, the recovery rate is the most important aspect for reducing total infection.

Moreover, if we compare the results from Simulation 1 with Simulation 2 (Figure 3.6), as well as the bifurcation analysis, we note that without some sort of cleaning rate, social response is unable to reduce the total infected population on its own unless its rate of response is higher than the rate of infection. This means that social responses without a systematic cleaning or recovery policy merely reduce the rate of spread, but not the overall infection equilibrium.

This offers an explanation for the continuing existence of, for example, the long past Blaster worm. [20]

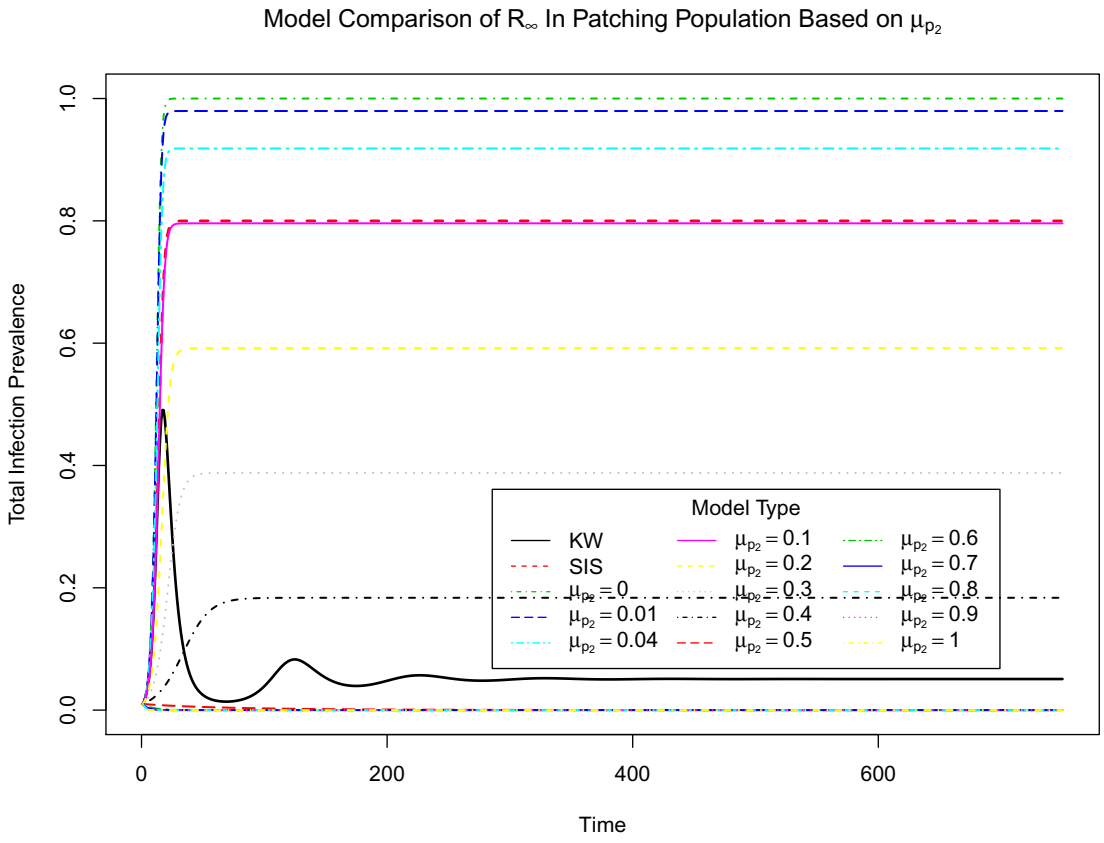


Figure 3.6: Comparison between SIS, KW, and Our Model, μ_{a_2} variable. Reductions in μ_{a_2} result in widespread prevalence. Moderate increases in μ_{a_2} can significantly reduce R_{∞_a}. When the rate of recovery exceeds β_a, $R_{\infty_a} = 0$.

Whether or not increasing the rate of social response to the level that it can be effective depends upon the topologies of the malware spread and social response. Enhancing social response is the focus of related human subjects research in progress. Exploring different malware diffusion and network topologies is the subject of future research.

While this model is not meant to model specific pieces of malware, but rather the prevalence of total malware, this result implies that there should be persistence of most previous forms of malware, just as the Blaster worm still propagates on older machines. [20] Under this model, and based on previous work, users of older machines are not willing, or able to identify and remove the older malware. Users with newer computers and operating systems have no impetus to search for ancient malware, leaving the older devices isolated.

This sort of isolation and persistence is very similar to the type of behavior we see in smoking cessation.

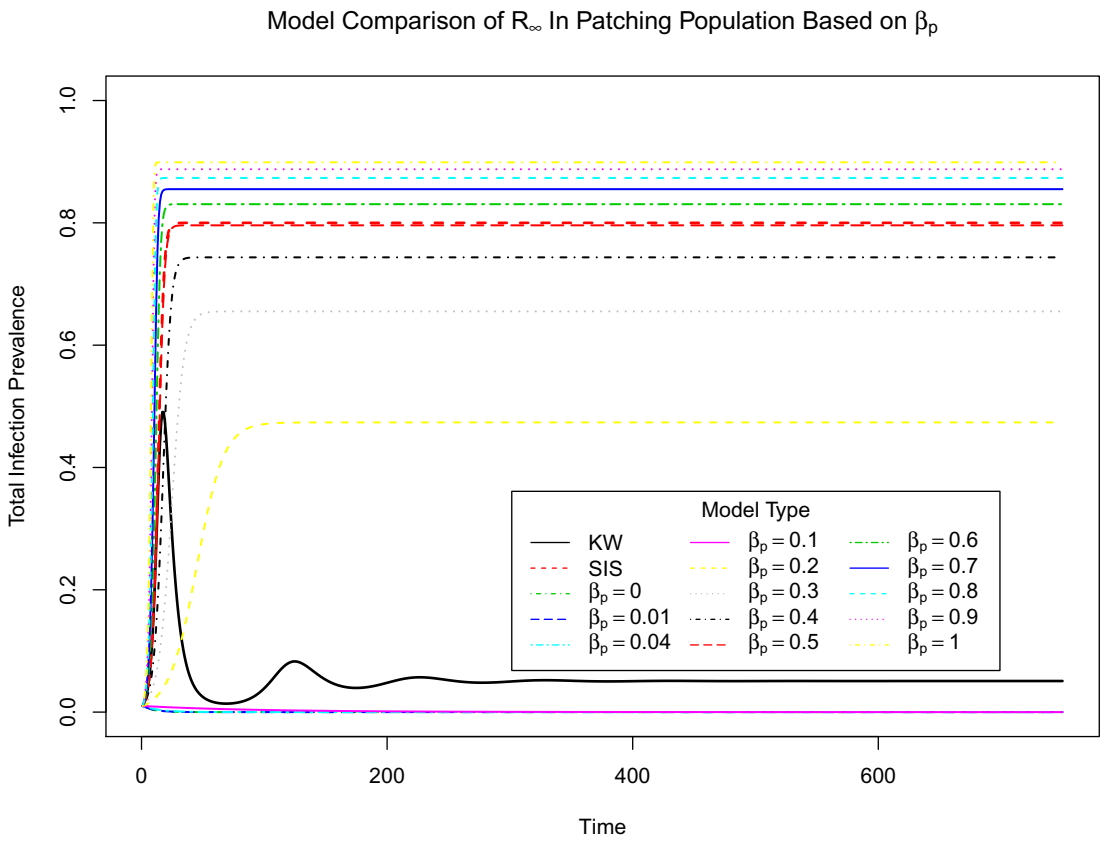


Figure 3.7: Comparison between SIS, KW, and Our Model, β_P variable. When β_a is reduced, global malware prevalence is reduced. Even at extreme values of β_a, μ_a is able to prevent R_{∞_a} from reaching 1.

The observed social effects of smoking cessation show that when individuals choose to stop smoking (μ in our model), if they are able to convince their neighbors to join them, they are successful. However, they are generally not able to convince all of their smoking neighbors to stop smoking (γ in our model). This leads to isolated pockets of smokers on the social network, and an endemic level of smoking. [47] This is exactly the type of behavior we see with older malware, with the isolated systems slowly being replaced with new devices that are no longer susceptible to the old malware.

Simulation 3

Simulation 3 is the complementary analysis to γ_{a_2}. There is no pandemic outbreak if $\beta_a - \gamma_{a_2} < \mu_{a_2}$. In practical terms, reducing β_a corresponds to user behavior, rather than cleaning/recovery of systems, as represented by μ_{a_2}. Examples of this include protected browsing, keeping a system up to date, or other security measures.

We see that, obviously, β_a is an important part of the infection. One thing that is noticable in our simulations is that β_a need not be very high to achieve a relatively high R_∞. When $\beta_a = 0.2$ (twice μ_{a_2}), $R_\infty = 1 - \frac{.1}{(0.2-.01)} = .474$, a value close to the 2 year average of infections provided by the APWG [10–12].

Effects of Adjusting A Single Parameters in Both Populations

For this set of analyses, we adjusted a single parameter in both vigilant and non-vigilant populations and investigated how it affected the R_∞ for the entire population. This allows us to study the global effect of a given parameter. The first three simulations examine the effects of behavior changes in the vigilant population. The final simulations study the interactions between the two uninfected populations governed by η and δ. We find that the principle way to reduce R_∞ is to allow infected individuals to recover to the vigilant population.

Simulation 4

In this simulation we adjust the social response parameter in the vigilant population. This allows us to see how increasing the parameters in vigilant population has on the system-wide R_∞ (Figure 3.8). We notice, as in the following two simulations, increasing the responses in the vigilant population does little to reduce the total R_∞.

The dynamic relationship between γ_{a_2} and R_∞ is a bit more complicated in this simulation, as this simulation contains an I_r value that is non-zero, and transitions between the populations. We are holding $\mu_{a_1} = 0$ and $\gamma_1 = 0$, so we know that the relationship between $\eta = 0.05$ and $\delta = 0.04$ gives us an approximate 80-20 split between security conscious users and those that are unable or unwilling to engage in more secure behaviors. We also know that when $\gamma_{a_2} + \mu_{a_2} > \beta_a$, $R_{\infty_a} = 0$, in an isolated situation.

However, even when $\gamma_{a_2} = 1$, we still end up with an infected vigilant population. In this case, $R_{\infty_a} \approx 0.072$, while $R_{\infty_r} \approx .6$. $R_{\infty_a} = 0.072$ represents approximately 31% of the vigilant population, while $R_{\infty_r} \approx .6$ is approximately 77% of the non-vigilant population. Recall Figure 3.5 that illustrated that with only a vigilant population $\gamma_{a_2} = 1$ should remove all contagion within the vigilant population. Thus, the infections within the vigilant population are being driven by the non-vigilant population.

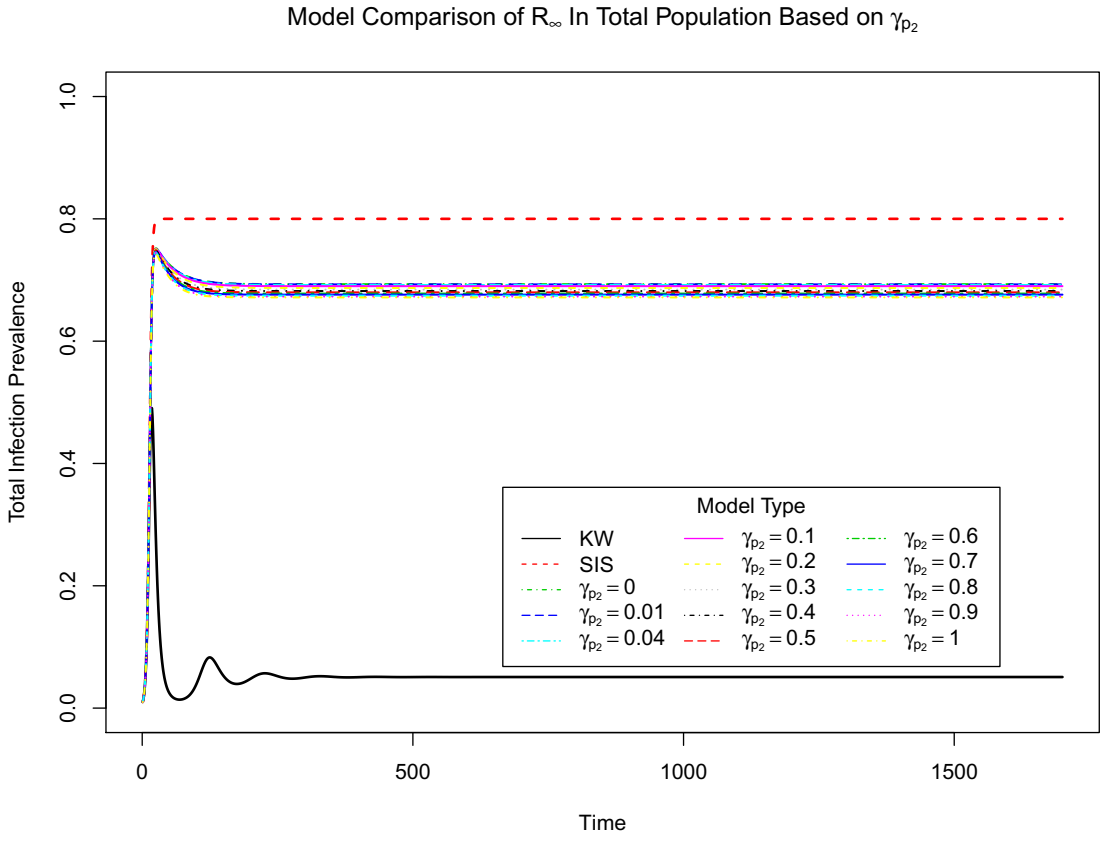


Figure 3.8: Comparison between SIS, KW, and Our Model, γ_{a_2} variable. Increasing the social response rate does reduce R_∞, but even at extreme values, γ_{a_2} is unable to significantly reduce R_∞ due to the effects of the large infected non-vigilant population.

Simulation 5

In this simulation we adjusted the cleaning rate within the vigilant population. The key result here is if $R_{\infty_a} > R_{\infty_r}$, R_{∞_a} drives the total R_∞ (Figure 3.9). However, this is unlikely, as it is improbable that vigilant users will become infected at a greater rate than non-vigilant users. While, if $R_{\infty_a} < R_{\infty_r}$, but $\gamma_{a_2} + \mu_{a_2} < \beta_a$, the infection is driven by both vigilant and non-vigilant populations. In the case where $R_{\infty_a} < R_{\infty_r}$, and $\gamma_{a_2} + \mu_{a_2} > \beta_a$, the infections in the security aware population are due to the prevalence of infectious non-vigilant systems.

For example, when $\mu_{a_2} = 0$, there is no cleaning and the social response cannot reduce the spread of the infection within the vigilant population. Thus, $R_{\infty_a} = 1$. At the end of the 1700 time steps in our simulation, $R_\infty = 1$, with most of it (99.998%) being made up of the "vigilant" population. This suggests that the

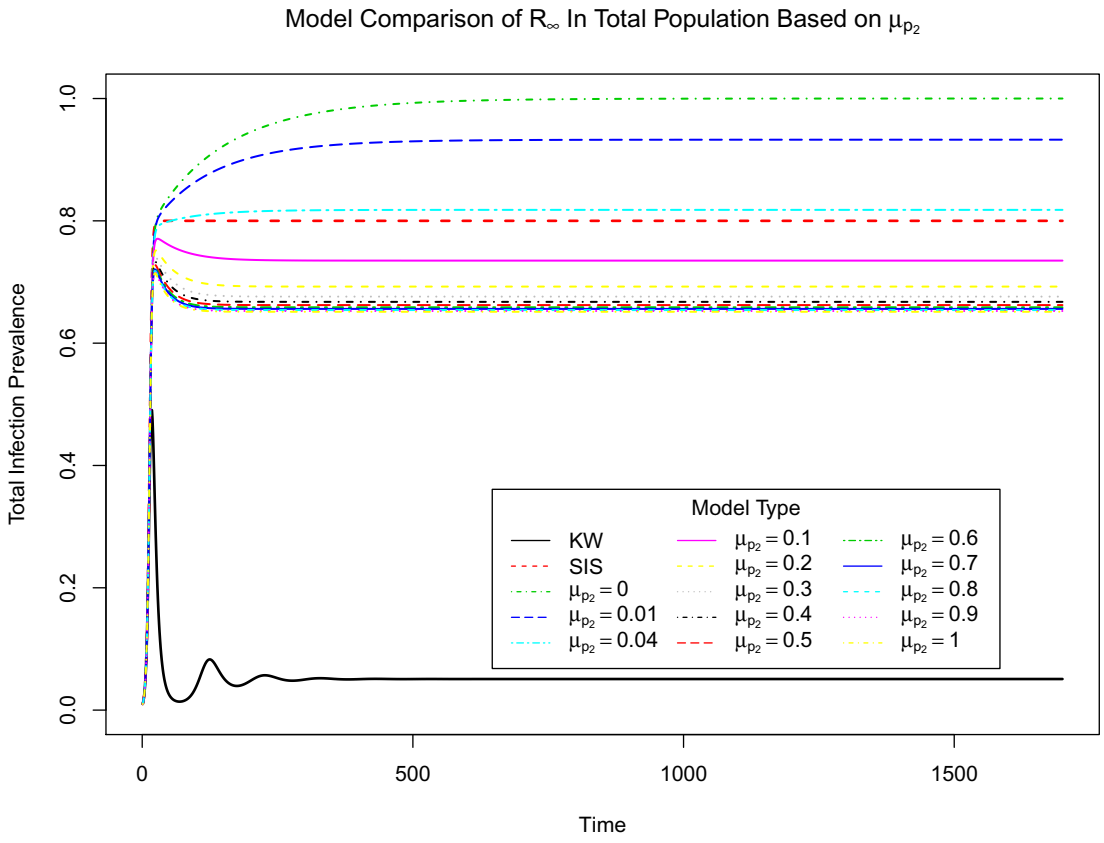


Figure 3.9: Comparison between SIS, KW, and Our Model, μ_{a_2} variable. Reductions in μ_{a_2} lead to increases in R_∞, as the global prevalence tends towards the behavior of the least secure population. Increases in μ_{a_2} are capable in reducing R_∞, but the reductions in R_∞ are mitigated by the behavior in the non-vigilant population.

vigilant population cannot rely merely on protective measures to avoid infection, but must also be diligent in actively monitoring and maintaining their systems.

Simulation 6

This simulation adjusted the β_a parameter to investigate how allowing the vigilant population to reduce, or increase its infection rate would affect the system-wide R_∞. Given the parameter values for μ_{a_2}, as β_a increases to 1, the vigilant populations dynamics approach those of the non-vigilant population. Hence the convergence to $R_\infty = 0.8$, as β_a goes to 1 (Figure 3.10).

What is interesting about this simulation is, if our recovery and social response parameters were such that the reproduction rate of the vigilant population were greater than the non-security aware population,

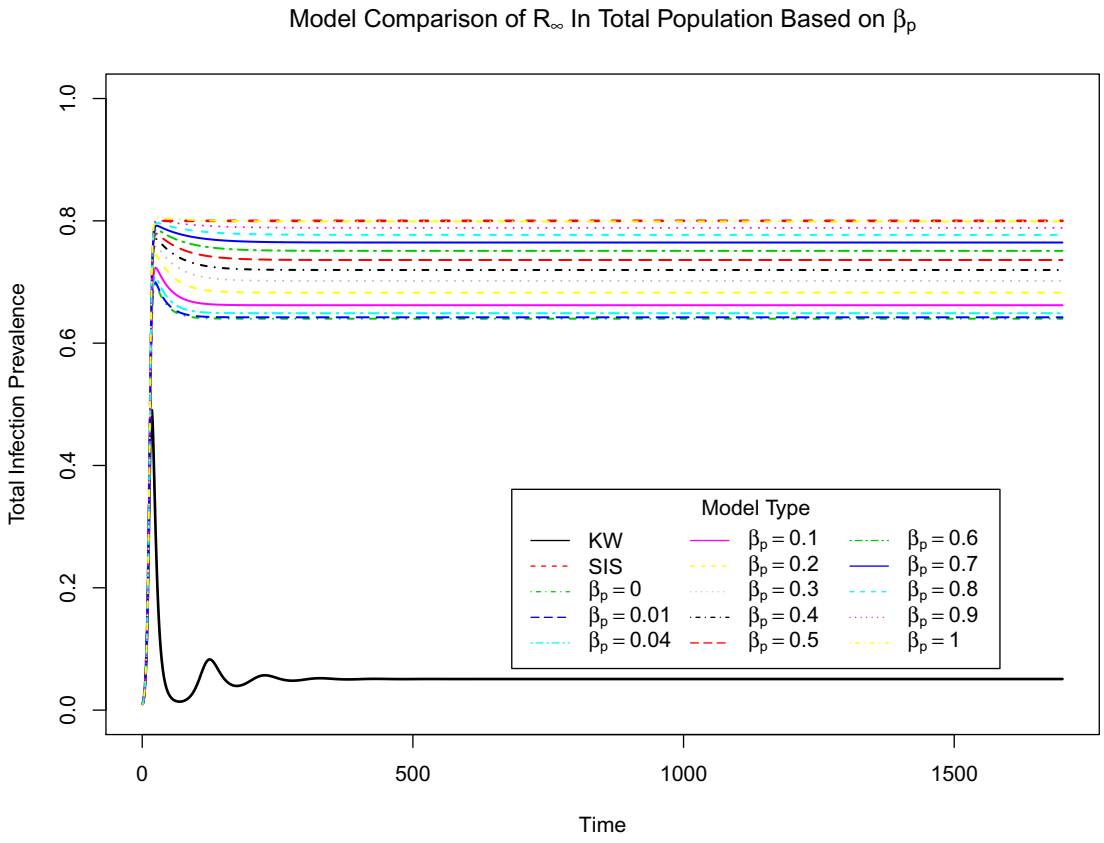


Figure 3.10: Comparison between SIS, KW, and Our Model, β_P variable. Increasing β_a increases R_∞, but due to the values of γ_{a_2} and μ_{a_2}, R_∞ performs as R_{∞_r}. If γ_2 or μ_{a_2} are reduced, R_∞ would also increase, as the vigilant population would be the least secure population, in that case.

it would pull the system to a total R_{∞_a}, as users would flee the infectious environment of the non-vigilant group, to the even more hostile vigilant group.

However, while $\beta_a > \beta_r$ is interesting, it is unlikely. It suggests that the risk averse population is more likely to be infected by a given piece of malware. In specific cases this might be the case, say where risk-averse users install infected security software and encourage risk-takers to install it. In this particular case, once the malware is discovered the cleaning response is likely to be more rapid, as the risk-averse population seeks to divest itself of the malicious software. However, in the limit, this is unlikely to happen with frequency due to increased effort and awareness necessary to maintain risk-averse behavior. However, it does suggest that, if this model is adapted to individual pieces of malware, that infection and social response should be tied to the expected utility of a given piece of software.

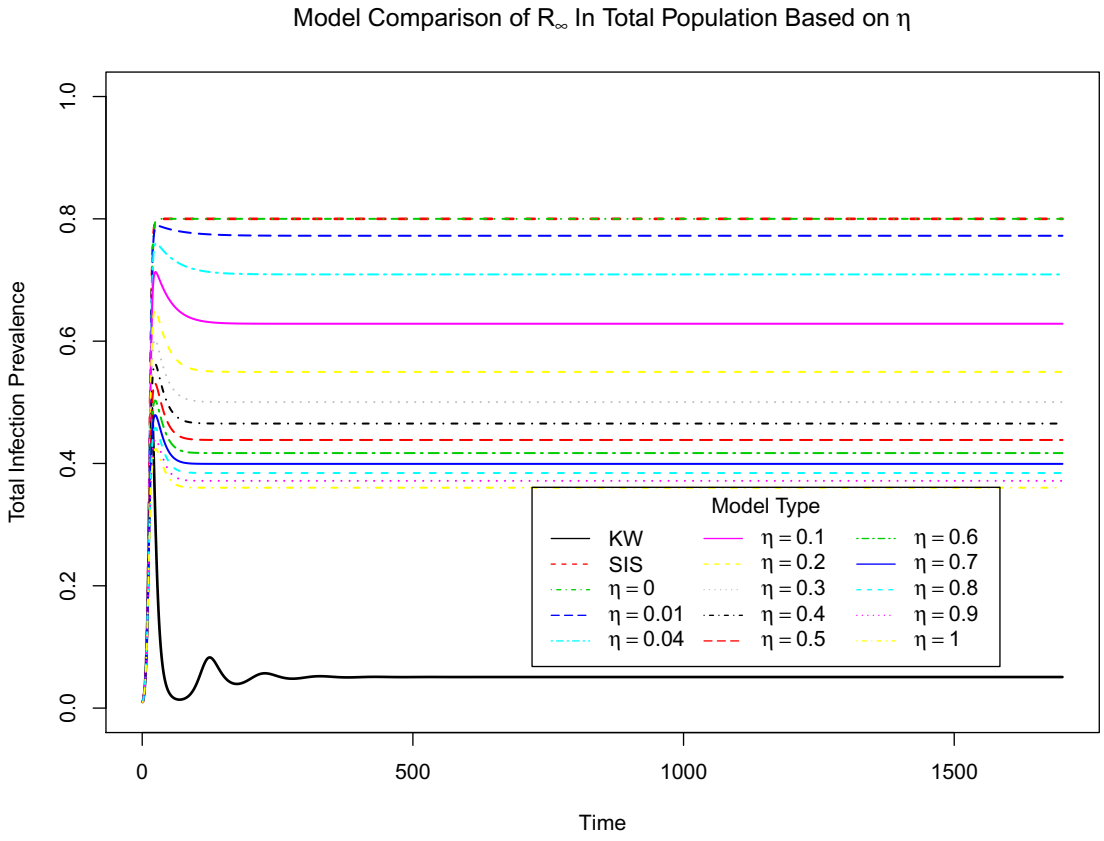


Figure 3.11: Comparison between SIS, KW, and Our Model, η variable. Increases in the ability for the susceptible non-vigilant population to become vigilant reduces R_∞.

Simulation 7

For this simulation, we varied η to see how increases in the response rate of non-security users in the face of infection impacted R_∞. Recall that η represents a user's ability to transition from non-vigilant, to vigilant, to reduce the likelihood of infection, in the face of an impending infection. Obviously, when $\eta = 0$, there is no transition to the security aware population, and the model behaves as a standard SIS model as shown in Figure 3.11.

However, when $\eta > 0$, the model behaves in an interesting manner. It is possible to see that increasing η reduces the system-wide R_∞ (Figure 3.11). But, there is a complex relationship going on between η and R_{∞_a}. As seen in Table 3.2 and Figure 3.11, while the total R_∞ is decreasing, the R_{∞_a} increases until $0.6 < \eta < 0.7$, when it begins to decrease. It is also possible to see that R_{∞_a}, while increasing in those intervals, is always decreasing as a percentage of the vigilant population.

Table 3.2: Interaction between η and R_{∞_a}

η	% Population$_a$	% Population$_a$ Infected	R_{∞_a}
0	0	-	0
0.1	40.1	43.4	.174
0.2	54.9	39.9	.219
0.3	63.1	37.5	.237
0.4	68.5	35.7	.245
0.5	72.3	34.3	.248
0.6	75.2	33.1	.249
0.7	77.4	32.1	.249
0.8	79.2	31.3	.248
0.9	80.8	30.5	.246

η pulls more of the total population into the vigilant population, but, until it is able to overcome the increasingly small non-vigilant population, that population still exerts a growing cost on the vigilant population. This result is important, since it indicates that even a small population engaged in risk behavior, with limited opportunity to reduce their risk, threatens a larger, risk averse population.

When $\eta \gg \delta$, it is unable to pull R_∞ to R_{∞_a} in the isolated system case. Yet even an η as low as 0.1, is capable of reducing R_∞ more than any of the test values of β_a, μ_{a_2}, or γ_{a_2}. This suggests that if modifying η is feasible, it would have a significant impact on global malware presence.

Evidence from information security and other fields suggest that increasing vigilance and awareness of risks is effective at reducing risky behaviors. [82] In particular, social networks are proving to be effective ways to disseminate information in a peer-to-peer like manner, specifically in regards to sexually transmitted infections. [99, 151] The reliance on social networks and awareness of affected neighbors is precisely what we encoded with η. While the effectiveness of risk communication on its own will never completely remove the prevalence of malware (Figure 3.11), reaching even a small proportion of the population can have noticeable effects on the propagation of malware.

Simulation 8

In this simulation we varied the other part of the transitions from non-vigilant to security. δ represents the constant rate of relapse where users view the costs of maintaining security may not be worth it. Reducing δ represents increasing a user's willingness to engage in more secure behavior, while increasing δ represents users that are only willing to be vigilant in the face of large outbreaks.

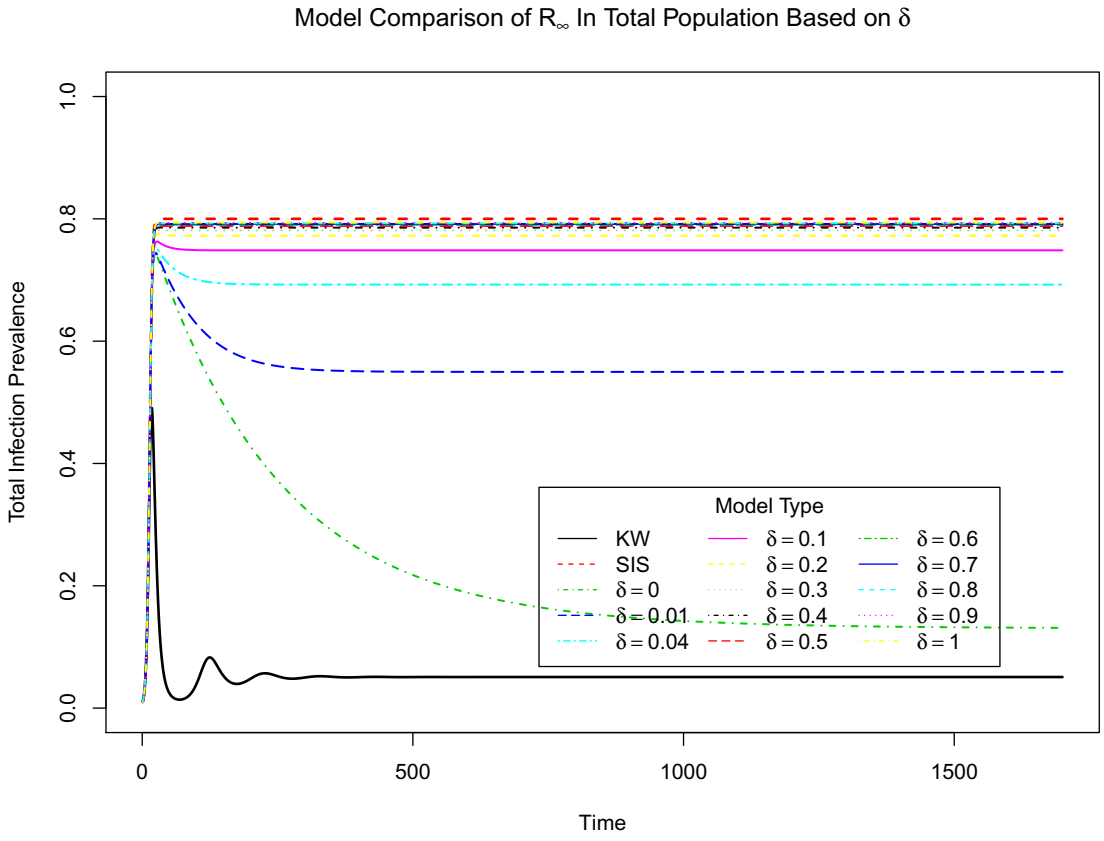


Figure 3.12: Comparison between SIS, KW, and Our Model, δ variable. Reducing users willingness to become non-vigilant reduces R_∞.

What becomes immediately apparent is that when $\delta = 0$, R_∞ approaches R_{∞_a} (Figure 3.12). However, $\delta = 0$, while ideal, is unlikely. It represents a population that is fully vigilant, irrespective of cost. We can see that reducing δ from 0.04 to 0.01, results in $R_\infty \approx .549$, which is lower than the R_∞ achievable by extreme values in β_a, μ_{a_2}, or γ_{a_2}. It is unlikely that such lack of sensitivity is realistically achievable, though it is probably reasonable to assume that η and δ are of the same order of magnitude.

Simulation 9

In this simulation we investigate the ability of users to recover to a vigilant population through social response, rather than merely recovering to the standard susceptible population. γ_{a_1} represents non-vigilant users' ability to respond to social pressure applied by non-infected vigilant users, not just to clean their machines, but to also, at least for some time, to become vigilant users.

In KW's social response model, γ_{a_1} is kept to 1/10 of standard cleaning rate, but is effective at reducing

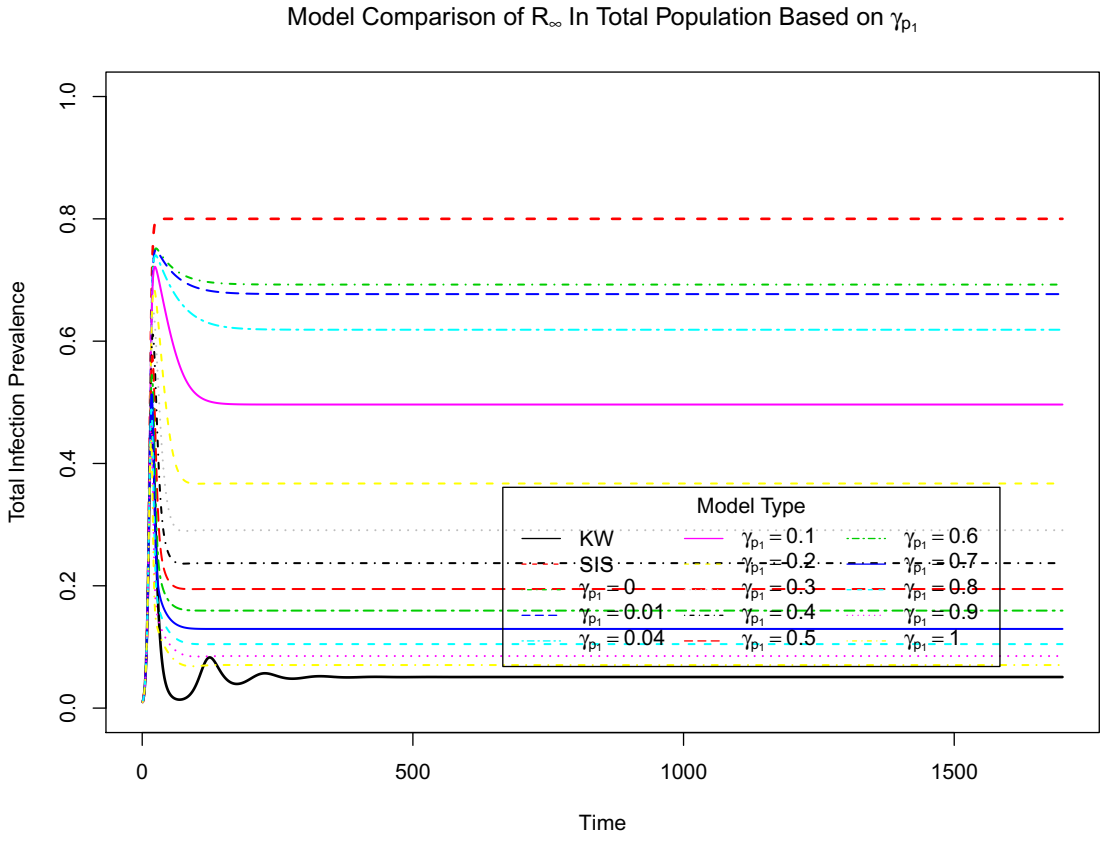


Figure 3.13: Comparison between SIS, KW, and Our Model, γ_{a_1} variable. Increases in a users' ability to become vigilant in response to social pressure is effective at reducing R_∞.

R_∞ due to the lack of infection rate in the recovered population, and the inability to recover directly back to the susceptible population [114]. We concur with their assumption, in terms of limiting the social response rate. However, it is important to note how effective increases in γ_{a_1} are at controlling width of the infection peak curve, and mitigating R_∞.

While the infection peak is not important for determining the total number of infected devices, it is useful to gauge the severity of an infection. High, narrow peaks indicate rapid infections, infecting most of the population before detection and cleaning. In particularly noxious infections, rapid peaks can be self-limiting, as overly ambitious malware chokes system and network resources trying to spread. [185]

On the other hand, slower curves may indicate longer lived propagation, with less of a given population at any given time, but eventually, potentially, just as many total infections. Obviously, if height of curve x over all time steps is less than that of another curve y, the total number of infections will be lower for x. [5, 104] There are slightly different behaviors in networked structures, particularly if there are multiple

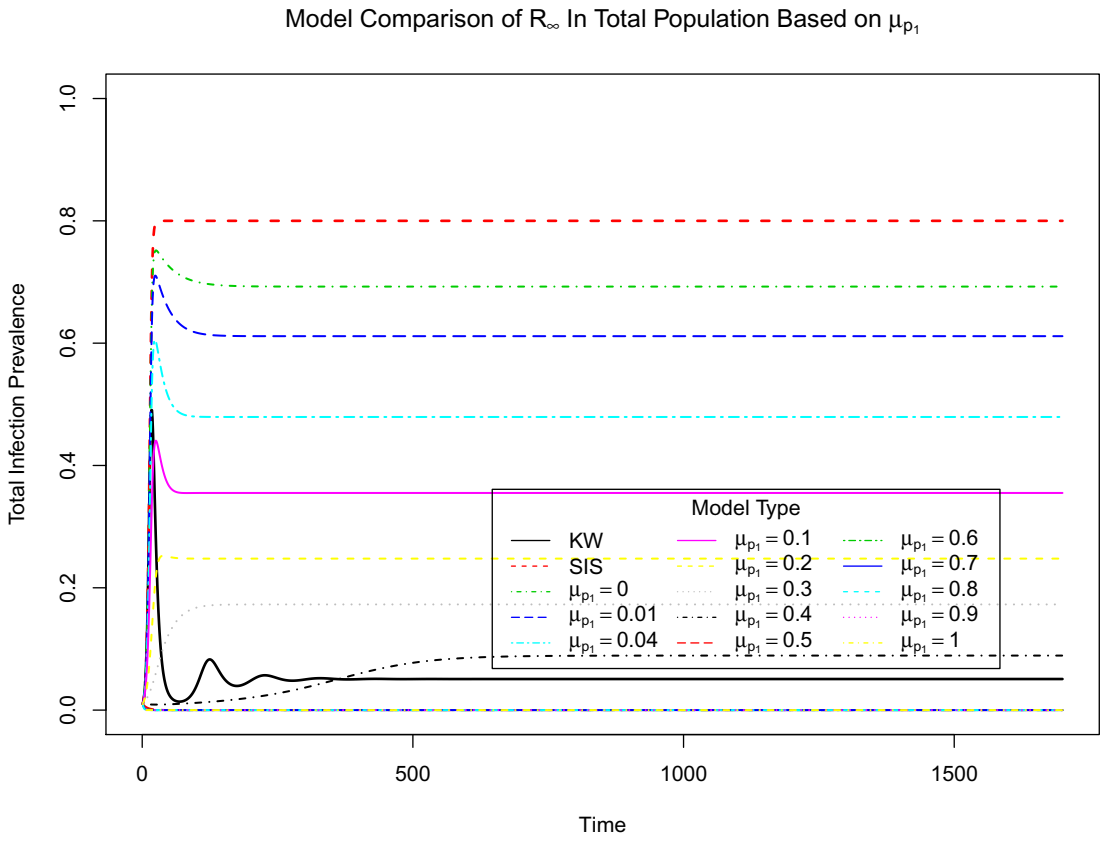


Figure 3.14: Comparison between SIS, KW, and Our Model, μ_{a_1} variable. Allowing users to recover their systems into the vigilant population is effective at reducing R_∞.

sources of a given infection, as we will discuss in Chapter 4. Arguably, γ_{a_1} should be limited in regards to β and μ, it may be, that given certain network topologies, even a relatively low γ_{a_1} will still be effective at reducing R_∞.

Simulation 10

In our final simulation, we vary μ_{a_1}, the parameter representing cleaning a computer and adapting vigilant behavior. For example, a user reinstalling an OS and applying patches and installing AV software, rather than just removing malware and hoping to avoid infection in the future. When $\mu_{a_1} = 0$, users are unable to become vigilant users until they clean their computers and respond to the infection through η. $\mu_{a_1} > 0$ means that users have some method to recover directly to vigilant behavior.

μ_{a_1} is not as effective as γ_{a_1} at limiting the duration of the infection peak, but it does limit the peak's height, limiting the total infections. Moreover, μ_{a_1} is effective at low parameter values. $\mu_{a_1} = 0.05$, or

1/10 β, reduces $R_\infty = .451$, nearly half the R_∞ of the standard SIS model, and without adjusting any other parameters. This suggests that providing users with the ability to recover to updated and secured software/machines, should be a key component in any campaign to limit global prevalence of malware.

Sensitivity Analysis

We used Latin hypercube sampling to examine the sensitivity of output variation to parameter variation [135]. The first step in LHS is to sample the parameter space to create a collection of measured outputs based on those samples. We did this sampling twice: first with all parameters sampled, followed by fixed values for the identified bifurcation parameters. In both cases we sampled the parameter space 1000 times. Our output of interest was total infection prevalence.

The results from the sensitivity analysis are interesting. Figure 3.15 shows the changing sensitivity of each parameter as time progresses. In the initial stages of the infection, social response and recovery from risk takers to the risk adverse population is more important than recovery within the risk averse population. However, it rapidly loses its importance on overall prevalence, while risk averse recovery increases its importance as time progresses.

All three of the standard recovery parameters (μ_x) are of approximately the same importance in the long term reduction of prevalence. However, the infection rate in the risk adverse group (β_P) loses its sensitivity gradually. The transmissions between susceptible risk takers and susceptible risk adverse (η and δ) are not significant in terms of affecting the global prevalence of a contagion, just as social response within the risk adverse community (γ_2).

When we fix the main bifurcation parameters ($\beta_r = 0.5$, $\beta_a = 0.25$, $\mu_{r_1} = 0.1$, $\mu_{a_1} = 0.01$, and $\mu_{a_2} = 0.02$), however, we get a better view of the effects of the social parameters.When all parameters are varied, γ_1 is significant parameter for reducing prevalence in the initial stages of a contagion, while γ_2 is never significant. However, when a contagion exists, we find that both of the social response recovery rates are important, at least until the later stages of a contagion (Figure 3.16).

Fitting the Model to Data

In our examination of the data we sorted each attack based on what online entity a given website was spoofing. We tabulated the total number of attacks on each entity to find the top ten targets of observed

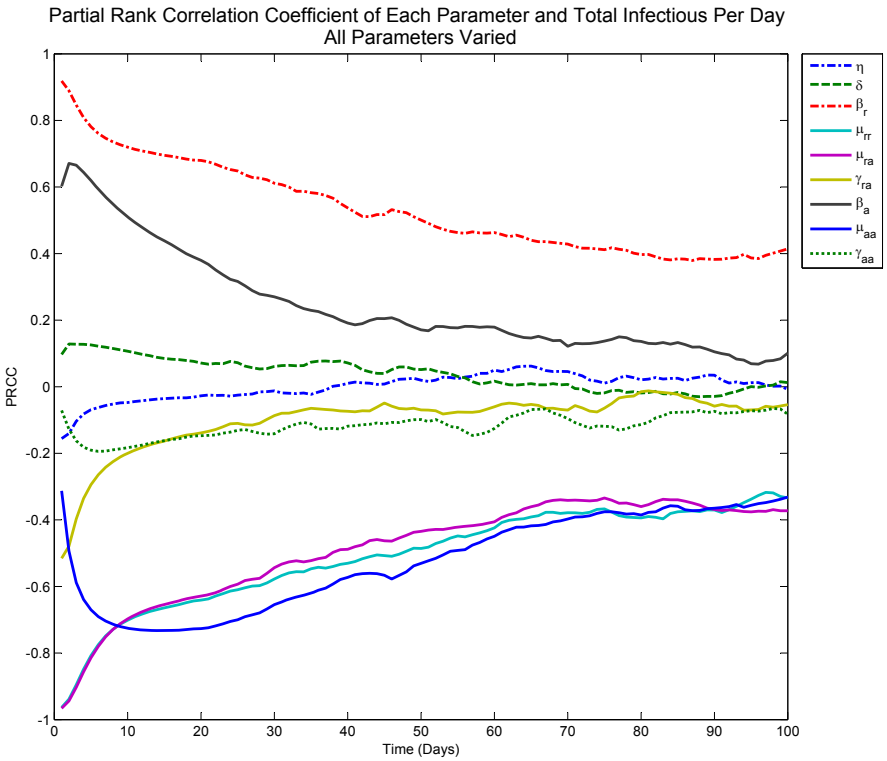


Figure 3.15: Partial ranked correlation coefficients for all parameters calculated for total infections at time=t.

attacks. These attacks have a heavy tail distribution (Figure 3.17, and seem to confirm Maillart and Sornette's work [132], but lower ranked entities still create large jitters in the overall behavior.

These jitters are expected, given the nature of the model. Consider that the model is not designed to model individual entities, nor individual malware. Rather, it is meant to model aggregate, global malware prevalence. In this regard, fitting the model against the most representative entity (Figure 3.18) and the total population (Figure **??**) makes sense. Fits against entities with fewer incidents, demonstrate the limitations of the model (Figures 3.22 and **??**), but also misrepresent the goal of the model.

We normalized the SSE to the minimum error across all fits and companies when we plotted the fits. Due to the complex interactions of the parameters, while the model can fit the data, often times the parameter values do not make sense. Thus, the sensitivity analyses is more important to this analysis, than the fit. However, the fits are useful to see how the model can represent real world data, and demonstrate the long term prevalence of attacks.

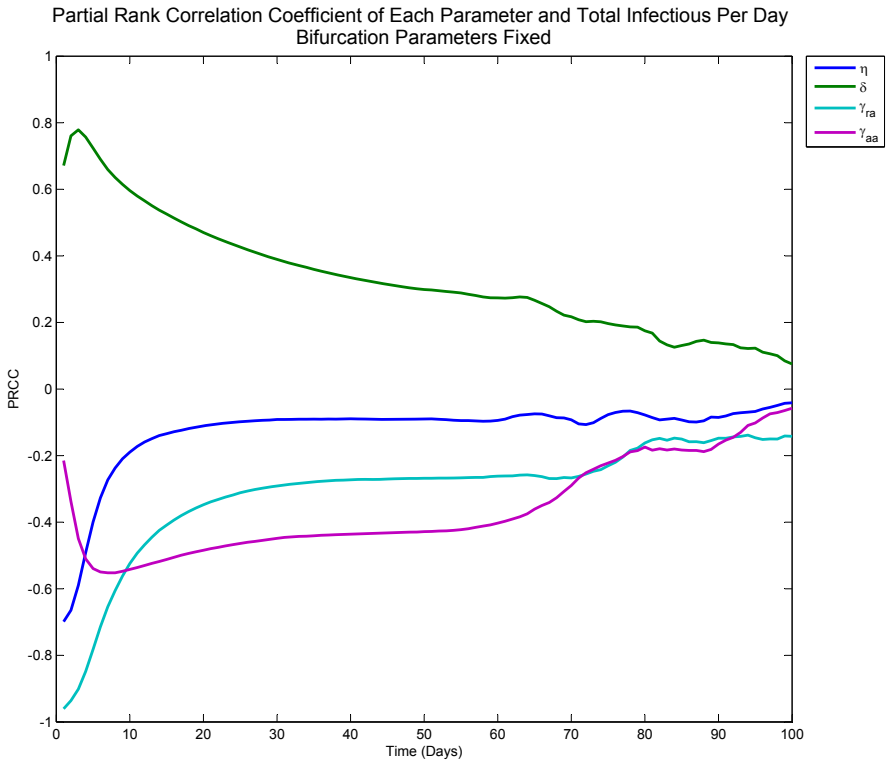


Figure 3.16: Partial ranked correlation coefficients for non-bifurcation parameters calculated for total infections at time=t.

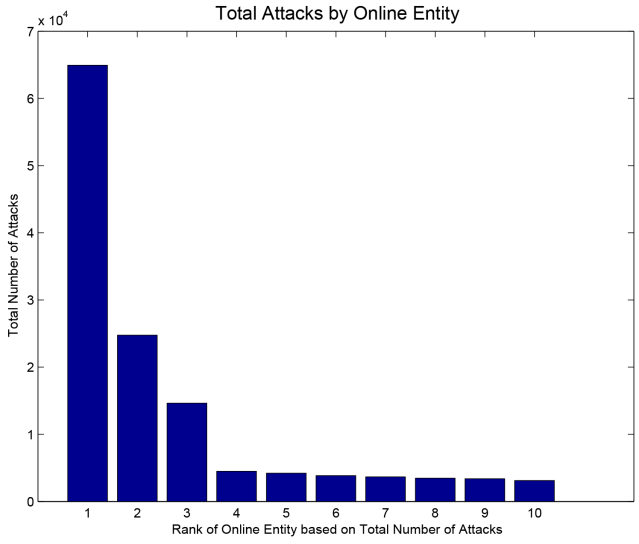


Figure 3.17: Plot of top-ten targeted entities by rank and total number of observed attacks.

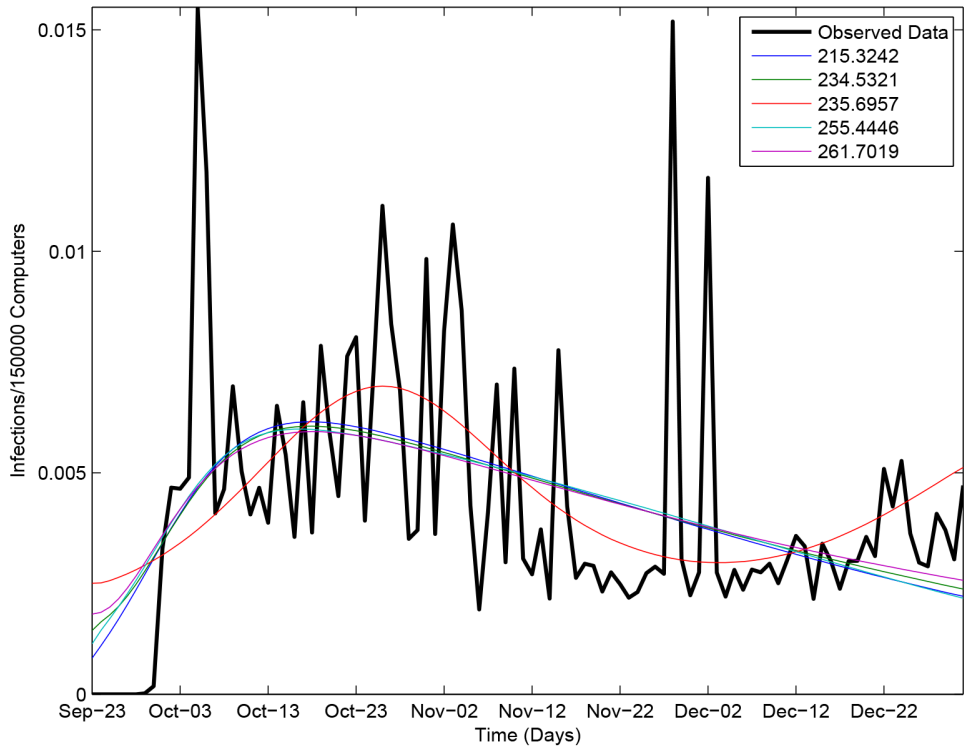


Figure 3.18: Top five model fits for the top targeted online entity

Model Fitting

While our model can handle the baseline prevalence of attacks in many cases (Figure 3.18), until we have additional data about risk-averseness, fitting this model with have limited use.This suggests that we need to have some further refinement to our model. Ideally, we would like to consider better representations of contact patterns, but also the exploration of birth/death rates, representing new computers entering the network and older computers being shut off.

Still, as a model for global prevalence, it seems to capture the important dynamics of awareness, risk-averseness and the costs of risk-averseness in aggregate. In particular, it can be used to conduct a mean field analysis of the top ten most infected countries (risk-takers) and the the top ten least infected countries (risk-averse) according to the APWG. The model suggests that collaboration between entities is necessary to implement more robust security and gives suggestions as to how to implement them, such as reducing the cost of risk-averse behavior.

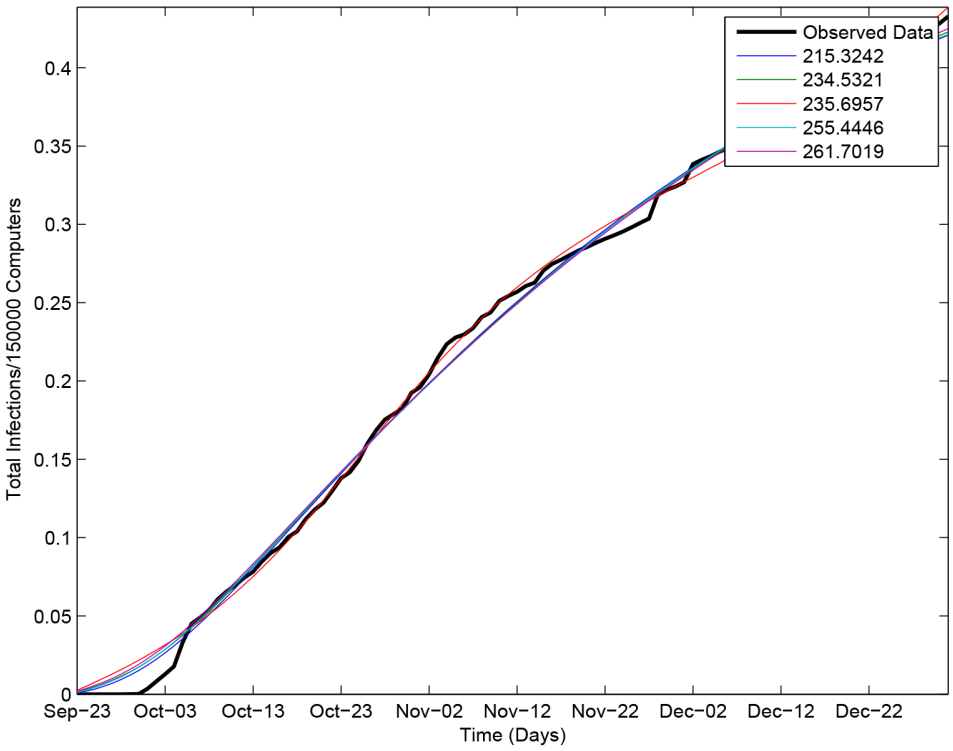


Figure 3.19: Top five model fits for the cumulative sum of observed attacks on the top targeted online entity

3.5 Discussion

In Section 3.6 we reify the conclusions of the ten simulations described in this work. In this section we discuss the possible implications of our findings. Extreme changes in β_a have little effect in the equilibrium state of an infection is an encouraging result. The rate of spread of an infection is one variable completely subject to the control of the attacker. Therefore great efficacy in changes in β would imply that defense could be ultimately futile.

Increasing the roughly equivalent variable, μ_a, is found to be as ineffective as β_a in decreasing the global prevalence of infection. However, there are a significant caveats. The outcome assumes that the malware will remain endemic with a roughly constant β and that recovery does not result in immunity to a particular malware component. Yet given the existence of multiple malware attacks, the use of multiple vectors for a single malware variant, the lack of broad immunity upon recovery, and the potential for malware to evolve these are not unreasonable assumptions.

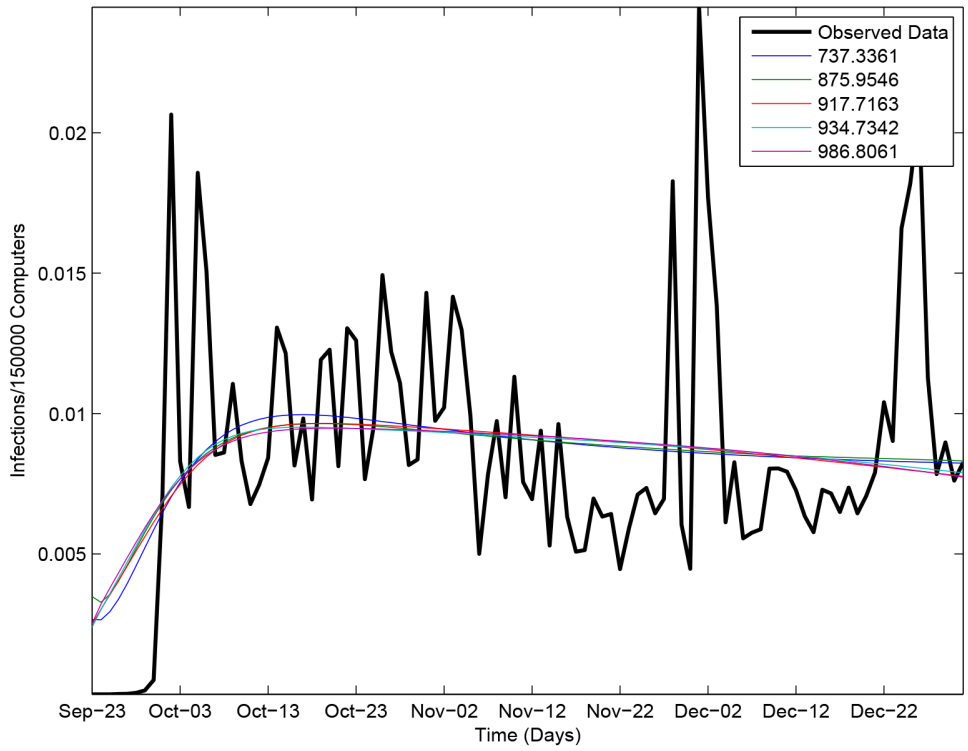


Figure 3.20: Top five model fits for the total attacks against the top ten targeted online entities.

Individuals choosing the recover due to social pressure (which includes automated pressure, such as Firefox exhortations to upgrade) must be faster than the rate at the virus is spreading. This is an extremely unlikely case. Yet the social recovery rate, γ_{a_1}, is one of the most effective measures in altering the equilibrium when there are two populations (vigilant and otherwise). However, increasing the response rate in the vigilant population has little effect on the global equilibrium. This is a mixed result given that it is arguably easier to alter a response rate in an aware population, but even modest gains in response of the unaware population can significantly reduce the global prevalence. However, since γ is a multiplier of μ, the combination of the two is very effective at driving down global prevalence.

Transfer rates between the two populations is the most efficacious strategy for reducing long-term equilibrium. This argues that small increases in vigilance can result in significant increases in outcomes just as increased use of healthy behaviors (e.g., contraception use or smoking cessation) can greatly reduce unintended consequences over the population as a whole. Compared this to situations where the entire population must engage in healthy behaviors (e.g., immunization) to result in significant outcomes. This argues for an

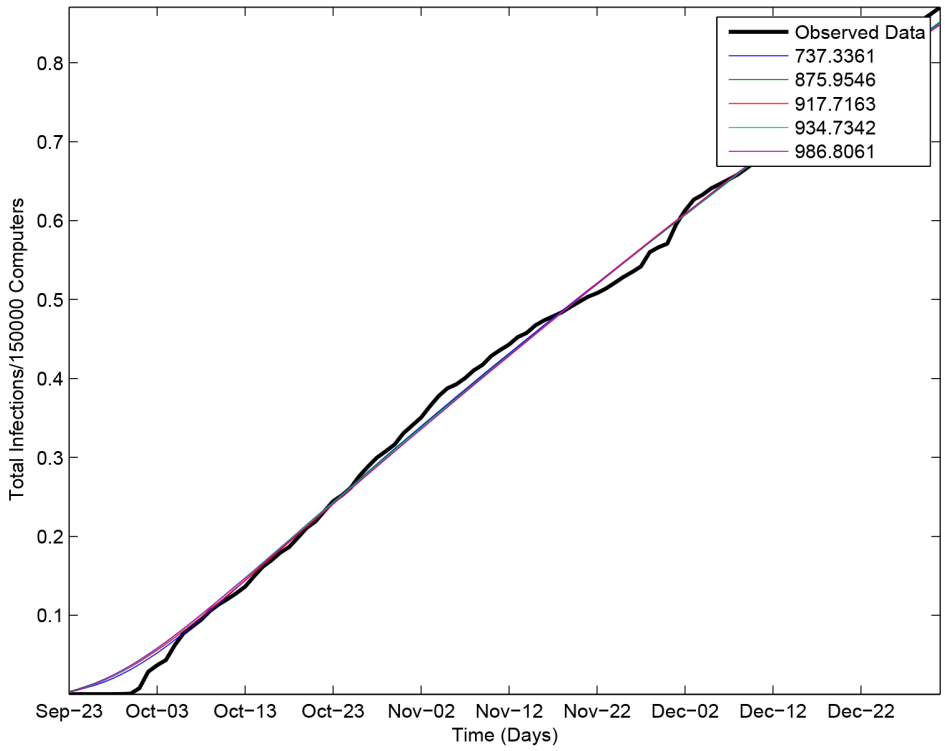


Figure 3.21: Top five model fits for the total cumulative sum of observed attacks on the top ten targeted online entities

approach that is closer to risk communication than mandates. Luckily, risk communication is feasible while global mandates are not.

Users must be able to act upon available information, e.g., δ should be quite low. The requires an ease of access to the resources necessary to engage in more secure behavior. Within the public health sector, barriers to treatment and preventative measures have been shown to greatly increase overall costs. For example, Franzini *et.al.*, estimated a likely additional cost of $43.6 million in a one year period in Texas, if adolescents were required to notify parents when they received reproductive health care [77]. This suggests that allowing access to security patches, even in the case of illegal copies, would be effective in lowering system-wide costs, though offering those patches may not be profit maximizing for a given firm [126].

Moreover, risk communication, when combined with access to treatment resources, has been effective in reducing prevalence in the public health sector. Spain *et.al.* demonstrated the effectiveness of at risk communication at recruiting at risk groups to utilize reproductive and preventative health care [195]. Several studies demonstrate the effectiveness of Youth Peer Education services at referring at risk populations

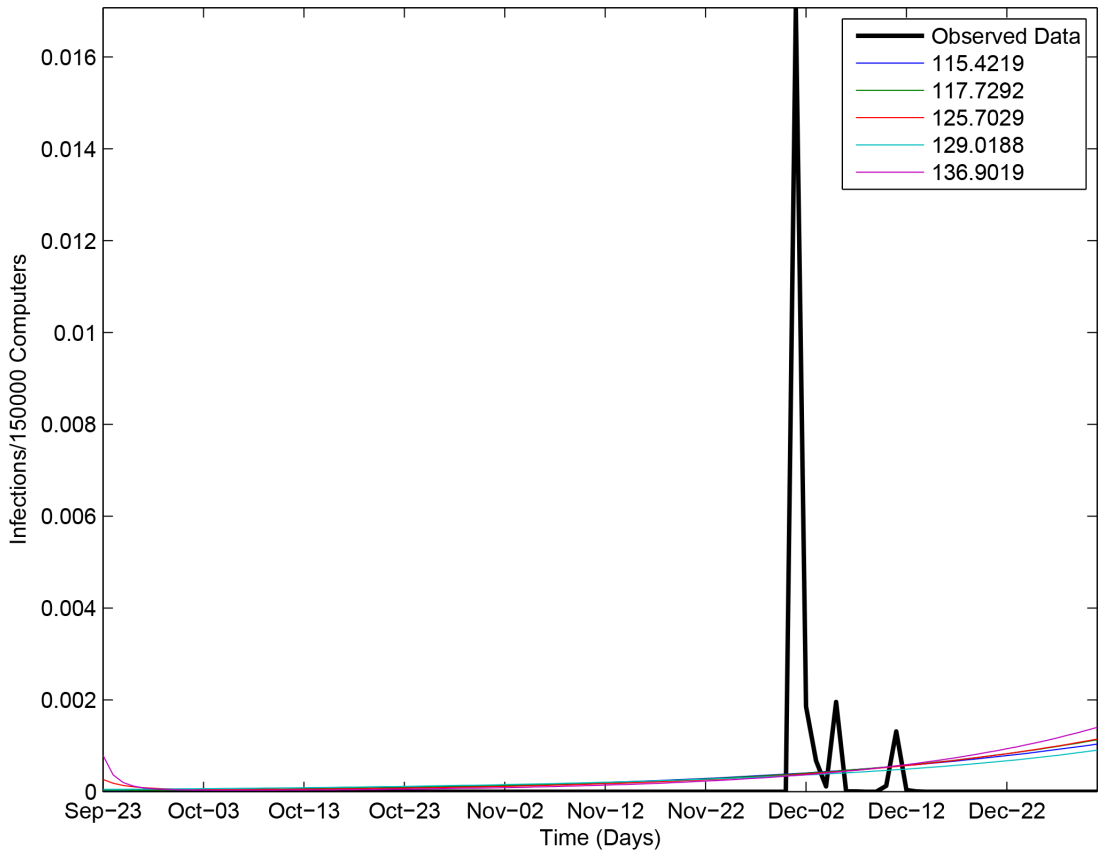


Figure 3.22: Top five model fits for the total attacks against the eighth most targeted online entity.

to appropriate clinics [37, 127]. When coupled with a voucher system for care, use of clinics increases dramatically [25]. Thus, there are extant systems of response and information that we can take advantage of in regards to encouraging more secure behavior.

Recalling the results of our sensitivity analysis, reducing δ is initially slightly less effective than risk communication between risk-averse population and risk-taking populations (γ_{ra}). However, the decline of the effectiveness of δ over time is less severe than the decline of γ_{ra}. To complicate matters, risk communication within the risk-averse population (γ_{aa}) is unimportant in the initial stages, but rapidly exceeds the effectiveness of both γ_{ra} and δ during the middle stages, but then fades rapidly, leaving γ_{ra} and δ roughly equivalent in terms of importance in the equilibrium portion of the endemic.

The problem with risk communication, however, is that finding effective measures of communicating risk is difficult, and, in terms of information security, a problem that has not been solved. There are some guidelines. For example, using video and relevant context to communicate risks to older populations. [82]. Training does seem to work to limit the effects of some forms of attack, but the effects of training do not seem

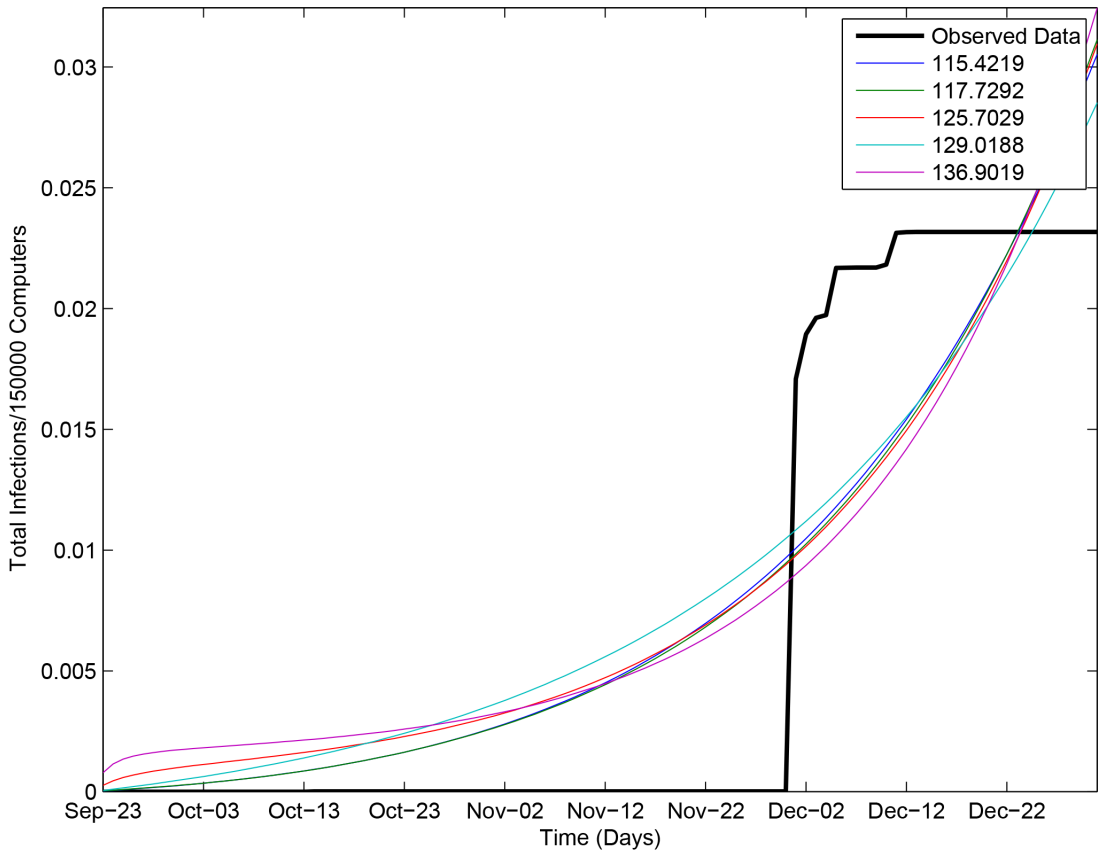


Figure 3.23: Top five model fits for the total cumulative sum of observed attacks on the eighth most targeted online entity

to persist long-term. [108, 177, 186] However, the specifics of risk communication in digital environments is a deep field that our model suggests can be effective.

3.6 Conclusions

In this chapter we created and examined the parameters of two-population SIS epidemiological model in regards to global prevalence of malware. The two populations, vigilant, and non-security aware, interact in many different ways (Figure 3.4), which affects R_∞, to equilibrium infected population. We examined single parameter variations within the vigilant population and the system as a whole to identify key components to addressing the spread of malware.

In our first set of simulations, we examined the vigilant population in isolation, seeking to identify the most effective parameter for reducing or removing malware in that population. We found that, within the single population it was possible to completely eliminate malware spread by setting $\mu_{aa} + \gamma_{aa} > \beta_{aa}$

(i.e., making the recovery rate and the social response greater than the infection rate). We also showed that adjusting the recovery rate μ_{aa} is the most effective way to reduce R_∞ in the vigilant population.

In our second set of simulations, we looked at the entire system and tried to find which parameters were effective at reducing global R_∞, while keeping the infection and recovery rates (β and μ) in the non-vigilant population constant. Here we find that, while we could eliminate the spread of infection within the vigilant population by overcoming the infection rate, adjusting the vigilant parameters had little effect on R_∞, and infections in the non-vigilant population drove the infections. However, when we examined the parameters governing transitions from non-vigilant to vigilant, we discovered several possibilities for infection control.

When we evaluated the transitions between uninfected non-vigilant and uninfected vigilant populations, we found that, while η was effective at making more users vigilant, even a small population of infected non-vigilant users could negatively impact the vigilant population. Similarly, when we prevented users from returning to the non-vigilant population, we could limit the infection spread to R_{∞_a}. Yet, this represents an unrealistic expectation of inelasticity (i.e., all users demanding secure behavior, regardless of cost).

Examining the parameters governing the recovery of infected non-vigilant to uninfected security aware, we find that allowing users to clean and repair their systems with updated and secure software is the most effective way at managing global prevalence of malware infection. Even at low levels, the ability to recover to the risk adverse population (μ_{a_1}) greatly reduces the global infection prevalence. Additionally, while not as effective as the risk adverse recovery rate (μ_{a_1}), the social response recovery to the risk adverse population (γ_{a_1}) is the next most effective parameter. This suggests that coupling social response, along with access to updates for all users, would be an effective measure for reducing the prevalence of global malware.

In the next chapter I will examine these effects of environment and population behavior in the context of mobile malware. The results from Chapter 4 demonstrate that social response, when taken in the context of social networks and mobile market places, is more effective than market place construction. In fact, even with low initial recovery rates, exploiting social responses can drastically reduce the spread of malware.

4 Connectivity and Privacy

- How might observed user preferences in mobile app adoption affect the spread of mobile malware spread via marketplaces?
- How do markets affect the spread of mobile malware?
- What is the risk, in terms of privacy loss, of mobile malware?

Our results from examining user behavior from the perspective of risk-averseness and social response led us to study the importance of social context and environment on malware spread. We designed a novel piece of mobile malware, as proof-of-concept, to capture contacts in participants' smart-phone address books. After testing its effectiveness in a small population pool (N = 24), we were able to capture over 4000 contacts. Malware matching the behavior of our proposed malware was discovered in the wild between our discovery and publication. [46, 235]

We then developed an individual-based model that captures the studied app adoption in Android and Apple users, as well as smart-phone application marketplace behavior. We then used our model to study how social pressure might drive the adoption on smart-phone applications, and how effective different marketplace configurations might be at controlling the spread of malware. We found that marketplaces do a good job in limiting the spread of mobile malware, assuming they are functioning properly. Users' application adoption strategies affect the spread of malware, but this is only noticable when they use non-vendor marketplaces. Finally, users' application adoption strategies seem to be well adapted to the markets they participate in.

4.1 Introduction

Cellular phones that utilize the Android, iOS, or Windows Phone operating systems (smart phones) represent a growing share of the personal computing market, making up roughly half of all cell phone users. [190]

With smart phones containing significant amounts of personal data, they have attracted the attention of malware developers and scammers. [72]

To address the security risks presented by the sensitive data stored by smart phones, Google relies on an explicit all-or-nothing permission scheme to limit the access Android apps have to a user's personal information. If a user accepts the required permissions the app is then installed and the permissions are granted to it. However, a user must accept all of the required permissions or the app will not be installed.

When a user attempts to install an app downloaded from the Google Play store, a list of permissions is shown and the user may tap or click on each required permission to see for what functionality the permission is required. When a user downloads an app from a 3rd-party source, such as a download from a website, however, the user is unable to expand the permission list to see what each actions each permission allows the app to perform.

Even in the case of the Google Play store, it is difficult to know whether or not an application will use the requested permissions in an approved manner. Thus, users rely on their social contacts and application popularity ratings, rather than security indicators, to make decisions when installing applications. [3,44,161]

Apple also claims that all apps are reviewed to meet Apple's quality expectations. In particular, Apple attempts to verify that apps are reliable and perform as expected. [14] Apple commonly rejects apps for misleading users and inaccurate descriptions of apps, and because Apple reviews the behavior of the app prior to deploying it on the app store, iOS users do not have to make decisions based on security permissions. [15] This may be one of the reasons that have a more balanced approach to downloading, relying on the total download counts and social networks, rather than focusing primarily on social networks, as found in Android users. [44]

4.2 Background

Application Installation

The current implementation of Android‘s permission system is inadequate. Despite users‘ demonstrated concern for privacy, many ignore the permission lists at the time of app installation. Technically complex terminology, interruption of the user's installation task, and the inability to compare permission lists between similar programs can prevent users from making informed decisions despite their demonstrated concern for privacy. [67]

A survey conducted by Chin et al. found participants are likely to install free apps from unknown developers and that the most popular methods of app discovery are searching app marketplaces and word-of-mouth, with recommendations from friends and family leading to increased downloads. [44] They found that while 44.8% of iPhone applications are adopted solely through search. Android users, on the other hand, adopt applications based on search in only 15.7% of instances, relying on recommendations from friends and family in a plurality of cases. [44] Particularly less technical users tend to seek advice from their social networks. [215]

In addition, research indicates that users depend more on advertisements and popularity, rather than security information, such as privacy policies, when they consider installing an application. [44] Work by Felt et al. suggests that this is due to the fact that the permission list that appears at install-time is explicit but difficult to understand. [73] They also suggest that users may be experiencing warning fatigue and do not want to be bothered to read popup notifications if the notifications interrupt their desired task, such as app installation. [71] More explicit or graphical warnings have mixed results. [26]

Warning fatigue is a well established phenomena accross the field of information security. Böhme and Köpsell found that even privacy and security aware users generally click through EULAs and privacy policies to accomplish their desired tasks. [30] Schechter et al. found that users generally ignore HTTPS indicators. [182] A more in depth look at users' ability and SSL warnings by Sunshine et al. found that in most cases, expert users were not different than non-experts, and most users were willing to ignore SSL warnings to get to the desired website. [202] In regards to mobile devices, Christin et al. were able to induce risk-aware users to install unknown mobile applications once the incentives were high enough. [48]

Aharony et al. found similar results by studying 130 adult members of a young-family living community for 15 months. [3] Aharony et al. studied Bluetooth proximity data of subjects, and found that those subjects that spend time together had more apps in common than those they spend little time with. Moreover, Pan et al. found that there were identifiable network effects in application installation, even with large amounts of uncertainty in overall installation behavior. [161] This research points to a social impact that can overcome initial concerns about unknown applications. [48, 161]

Malware Threat

Despite Google's assertions that malware would not be a problem on Android devices, there have been several instances of outbreaks. [52, 179] Even with Google's ability to remove apps from the store when

reported by users and its ability to remotely remove offending apps from users' phones, it is fairly easy for malware to spread to thousands of phones before it is removed. [72] For example, it was reported that the *DroidDream* malware was able to infect 260,000 devices in 48 hours before Google removed the offending app from the market. Of course, if we consider 500 million Android devices, the number of devices claimed by Erick Schmidt, CEO of Google, this is an infection rate of only .052 per 100 devices. [38] In addition, malware quite similar to our own was found on the Apple store shortly after we developed our application. [231]

While Google states that when uploading to the Play Store an app must meet certain criteria, which includes restricting apps from sending SMS or e-mail without user interaction, it has been reported that the company does not review apps before being added to the store. [44, 72, 87] Google removes apps from the store that are reported and confirmed as malicious, however, and the company has the ability to remotely remove offending apps from users' phones. [72]. However, alternative marketplaces can continue to host a malicious app even after Google removes it from the Play Store (Figure 4.1). [238]

Zhou et al. studied the prevalence of malware apps on different Android marketplaces. Of the 204,040 apps they downloaded, they found 211 malware infected apps, roughly 0.1%. However, the official marketplace had an infection rate of 0.02%, while alternative marketplaces had much higher infection rates. [238] Another study, conducted by Enck et al. found no malware in 1,100 applications, but found rampant misuse of privacy settings. [69] These studies show, that while currently there is little malware on official marketplaces, the sensitivity of the information found on smart phones and the centralized nature of the marketplaces makes even this small rate of infection a threat. [238]

Detecting mobile malware with technical means is also a problem. Enck et al. suggest using sets of permission requests as part of a strategy to detect mobile malware. However, Zhou et al. show that in many cases malware has a smaller permission footprint than legitimate apps and legitimate apps often violate the permission rulesets suggested by Enck et al. [70, 72] Pilz and Schindler studied the effectiveness of virus scanners for mobile devices and found that the best scanners identified only 50% of the most common mobile malware. [169]

Moreover, the virus scanners required a significant number of permissions and would be considered dangerous applications using Enck et al.'s permission sets. [70, 169] For example, Kaspersky Mobile Security violates rules 6 and 7 in Enck et al.'s ruleset by requiring permissions to receive, send, and write SMS messages. [169]

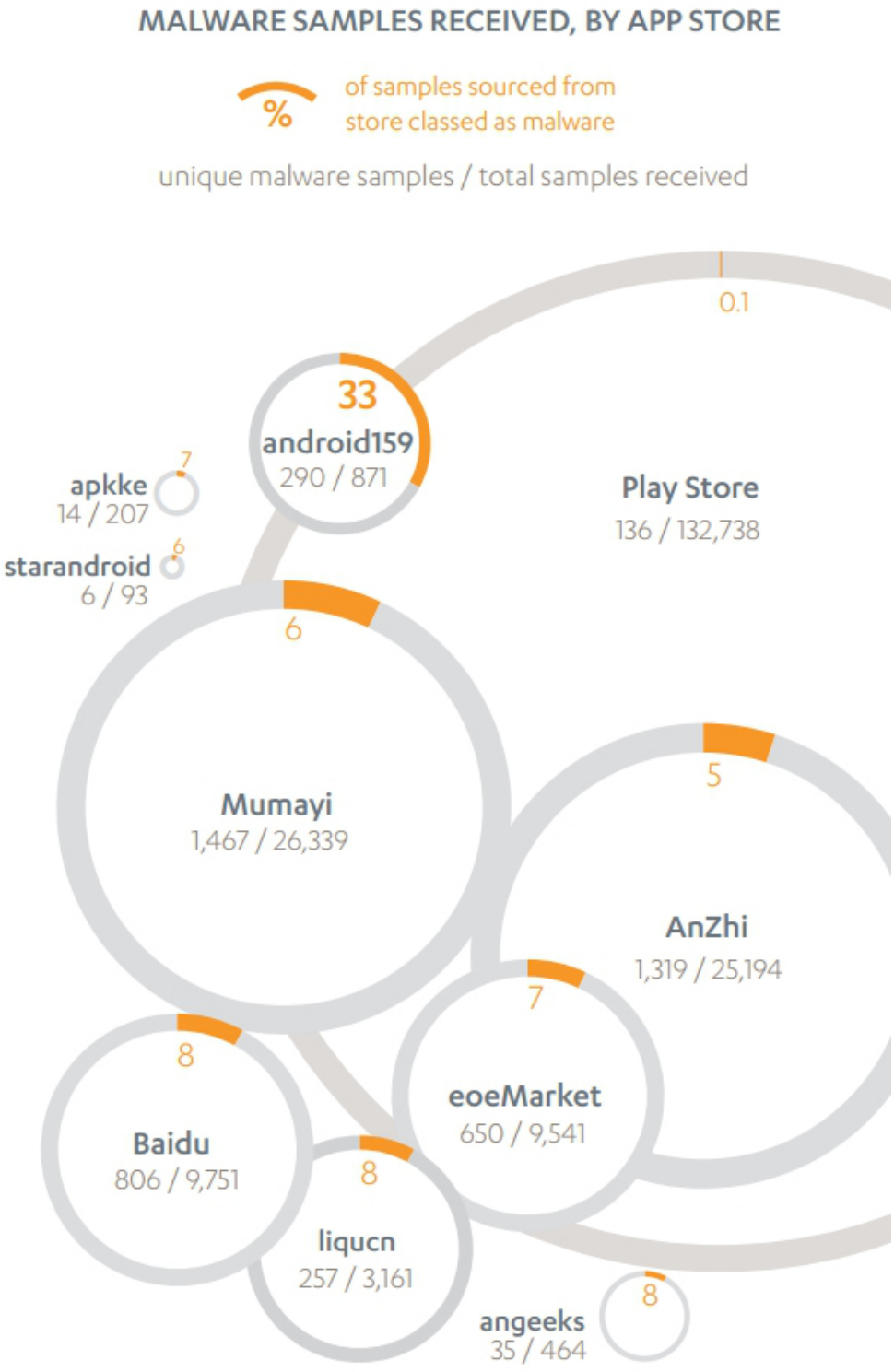


Figure 4.1: F-Secure Analysis of Malicious Apps on Given App Stores. Taken from [192]

Users' satisficing in terms of application installation and their reliance on social cues, rather than security indicators, creates an exploitable social vulnerability, with potentially large gains. Felt et al. examine a fairly comprehensive list of economic incentives implemented in mobile malware. [72] The most common of these incentives target stealing of credentials and personal information and the use of premium rate calls and SMS messages. There is an active marketplace for valid email and contact information, making the data exfiltrated by our experimental malware easy to monetize.

Modeling Infection Spread

Modeling malware spread has a long history starting with Kephart and White's work modeling the spread of viruses on directed networks. [113] As discussed in Chapter 3.2 their first and subsequent work on social and organizational recovery is concerned with the effect of network topology on malware spread and its use to represent realistic interactions. [114] Their work provided important theoretical insights in terms of the effects of network topology; however, they used random networks rather than grounding their work in social networks.

At the same time, with the emergence of HIV/AIDS, standard mass action assumptions of proportional mixing and homogeneous populations in epidemiological models were proving ineffective at describing the differences between in-group and out-group interactions. [103] For example, assuming statistical independence in contacts leads to bizarre situations such as, "female prostitutes and highly active male homosexuals, by virtue of their high relative activity levels, [having] a great deal of interaction" (i.e., model results could be explained only if male homosexuals were extremely active with female prostitutes). [148] When there are heterogenous populations and diverse social networks, simple differential equations will not suffice.

The use of network analysis as a way to represent non-homogeneous connections between people began to catch on within the larger epidemiological community shortly after with Pastor-Satorras and Vespignani's work on the spread of infections on scale-free networks. [164] Most of the modeling of computer malware focused on worms such as Code Red–which behaved similarly to random, or homogeneous, networks–examining more realistic models of infection due to timing and communication limitations rather than network structure. [218, 240] Still, the use of network analysis allowed for more nuanced approaches to the spread of viruses on networks like email connections and subnetwork topologies. [41, 80, 157]

The unification of the two strategies came in the form of attempting to model effective mitigation techniques such as those proposed by Staniford, Paxson, and Weaver. [196] In particular, modeling techniques

have been applied to methods for identifying outbreaks and methods for mitigating infection spread. Wang, Knight, and Elder studied the effects of network topology on immunization strategies, while other studies focused on the requirements of successful malware quarantine. [118, 144]

More recent work on mitigation has focused on social interactions and cognitive responses to knowledge of infection, rather than relying solely on centralized control. Perra et al. propose a general framework for behavioral models which allow for agents to change strategy in response to perception of infections, and Sahneh and Scoglio use a similar model structure, but use discrete models on network structures to demonstrate the effectiveness of self-quarantine. [55, 168] Stephan and Liò propose a similar framework that incorporates risk-perception in strategy selection. [117] Kelley and Camp further developed these models to examine risky and risk-averse behaviors for online interactions. [112]

The current models of mobile malware focus on heterogeneous connections due to agent movement patterns. Mickens and Noble use a random waypoint network for determining connections between mobile devices and model the spread of infection as agents move through their simulated environment. [139] Husted and Myers provide a model of mobile-to-mobile wireless malware spread including both timing parameters for infection and data-driven mobility simulation data. [105] Szongott, Henne, and Smith utilize a similar framework to study the spread of an iOS malware through multiple different population and location behaviors. [203] Wang et al. investigate patching behavior that affects mobile operating systems, but no one to our knowledge has studied the effects of subtle malware and social influence. [217]

We build on current research on the need to reconsider the current design of the Android permission system by developing a malware application that exploits users' consent. We also explore possible economic ramifications of this vulnerability by developing a model of malware spread and track information exfiltration as the malware is spread through a social network. We use our model to compare security effects for different market place strategies as well as strategies for maximizing the economic output of infected devices.

4.3 Design

Our study consisted of two portions, an experimental test phase and a modeling phase. In the experimental phase, we developed a malware application for the Android OS. We were unable to install the malware on participants' devices but were able to simulate infection. This simulation acquired the subject's email

contacts to give us a rudimentary estimate of the effectiveness of acquiring contact information.

To demonstrate the effectiveness of mobile-device malware, we developed a score-keeping malware app on the Android OS with simple contact exfiltration and solicitation of contacts to install the score-keeping malware themselves. When installed, the malware collects a subject's address book data and then sends SMS messages exorting contacts to adopt said malware. Using this application we reconstructed one hop of a social network from email contacts from 24 subjects that participated in our experiment on unknown app installation.

We then built an individual-based epidemological model with network interactions informed by MITs friends & family dataset. [3] We modeled individual behavior using previous human subjects experimentation in application adoption and installation. [44] The behavior of malware in our model was based on the execution of our own proof-of-concept malware, as well as limited information from news reports describing number of users infected within a given time frame. [179]

This individual-based model differs from the model in Chapter 3 by explicitly tracking individual infections using stochastic processes. We examined two different copying strategies. The first is based on preference for social recommendations; the second is a balanced approach, slightly favoring attention to global download counts. These two behaviors are associated with observed preferences in Android and Apple mobile device users. When we talk about user behavior, we are talking about these observed preferences in app adoption strategy.

We simulate an individual's behavior by drawing a random number each time step. If that random number is under a particular threshold, the individual updates its state based on its current state. There are three thresholds for infection based on an individual's preference to download an app based on random selection, attention to global download counts, and preference for local recommendations. Each individual has the same chance to randomly select malware. Each individual has the same chance to adopt an app based on global download counts, but this is, of course, scaled by downloads. Similarly, the reliance on local recommendations is the same for each individual, but it is adjusted by number of contacts and the amount of communication between those contacts.

As part of the model, we also designed market place behavior on documented market terms of service and academic and industry research into number of applications available on a given market place. [88, 238] In particular, we tried to implement the way centralized markets can affect installed applications. We also tried to implement accurate approximations of market place size based on available apps. [238] We also

tried to model market place responsiveness based on the size of the market place and amount of malware found on extant markets (Figure 4.1).

Our model illustrates the spread of mobile-device malware on a social network. Due to the limited size of our collected data, we constructed a weighted, directed network based on contacts between subjects in the Friends and Family data set. We then evaluated the diffusion of the socially-spread malware, the loss of privacy due to the spread of the malware, and the effect marketplaces have on the spread of mobile-device malware.

Our model demonstrates that a responsive official marketplace can effectively mitigate malware spread, but that the connected nature of social networks makes information exfiltration difficult to contain. The results of our model simulations illustrates the threat of socially-enabled malware diffusion. It suggests that there is an economic incentive to develop mobile-device malware. Current protections against loss of data, such as the permission system, are inadequate. The results of our simulations also show that there are numerous factors with the potential to limit the spread of mobile-device malware, including user behaviors and marketplace responsiveness. However, permission transparency does not appear to be an effective counter-measure.

Experimental Design

Our experiment consisted of two phases. The first of these phases was the design and testing of the trojan horse app. The second phase consisted of a survey of 24 students from a major university located in the United States. This survey collected simple demographic data and the collection of the user's contact list from their smart phones.

The design of the trojan horse has three primary components:

- The user experience
- The malicious code
- A simple Java server set to collect the information using sockets.

The user experience is an app that helps the user maintain scores for a variety of card and board games. This app was somewhat simple to create and apps of this type have approximately 50,000 downloads on the

Google Play store. The amount of downloads show that this type of app is useful, that people will download it, and provides a baseline from which to begin our modeling calculations.

When the app loads, the user is greeted with a list of games for which she may keep score. As she progresses through the UI, the malicious code runs in a background thread and steals her contacts. The malicious code operates in a secondary thread so that network activity doesn't impact the user experience, even in the case of communication error. The malicious code performs three actions:

- Access the contact list and search for contacts with e-mail addresses
- Stores each contact and information as an element in an array
- Transmits the completed array via plaintext to our server via sockets

The data collected about the user's contacts during our experiment was limited to e-mail addresses. Other data that can be collected include name, phone number(s), physical address, photos, and what group the contact belongs to in the user's Gmail account. However, the primary information we are concerned with is the contact's e-mail address. We hashed each e-mail address to maintain privacy while allowing us to compare addresses for uniqueness.

This attack is successful because the user does not notice the collection and transmission of data while the app runs. The user experience does not rely on network communications to complete desired tasks. The sole purpose of network communications in this case is to transmit the collected data and send it to a remote server.

Limitations

The malware app we wrote does have its limitations. The most important limitation is that the user is required to grant all of the required permissions to it. The permissions required for this app to run are:

- Read from the contact list
- Internet (to send the contacts over the network)
- Send SMS (for the SMS spam)

If the user notices the permissions and questions the legitimacy of the app based on the permissions requested, namely the permissions required to send SMS messages or access the network, she may not install

the app. It should be noted that the 'Internet' permission, required to send the contact information over network sockets, is requested by a great number of apps and is unlikely to attract much attention. [238] We also do not believe this is a major limitation because research suggests that people tend to ignore interruptive dialogs and will simply click through them to achieve the desired task; in this case installing the app. [30]

If the user exits the app before the contacts are sent then nothing malicious will happen. This can be changed by simply sending each contact separately instead of sending the entire contacts array at one time. However, if this approach is taken one could notice a repeated pattern of small data packets sent in regular increments through a period of time, thus possibly displaying a traffic signature.

It should also be noted that the sending of the 200 contacts during testing took approximately 8 seconds from the time the first contact was added to the array to the time the server received the completed array. Some timing differences will exist due to different connection speeds because of Wi-Fi, 3G, or 4G access. For the average smartphone user, this should not be problematic. However, for users with very large contact lists and slow speeds, this could pose some challenges in collecting the user's contacts.

Another limitation to be noted would be that the app's malicious functionality would be stopped if the user's phone is not connected to the Internet. However, as the majority of smart phones are connected the majority of the time we don't believe this to be a serious limitation.

Model Creation

Users' reliance on friends and family for app adoption, combined with the heirarchical interactions between marketplaces and devices presents an interesting modeling problem. [44, 161] Within mobile devices, there is both spontaneous and popularity driven adoption of applications. The popularity effect can be further divided. We divide it into a social network effect and a global popularity effect. While apps can be installed independently, the largest vector are marketplaces hosting malware, thus infections are facilitated by infectious marketplaces. On the other hand, responsive marketplaces can quickly remove malicious applications from their catalogs and greatly reduce the spread of a given malicious application. However, a user's desire to use an application may compel them to use third-party marketplaces to reinstall an app they may not recognize as malicious.

Model Assumptions

To model the behavior of mobile devices, we used a modified Susceptible-Infected-Recovered (SIR) model. The modified model allowed mobile devices to transition backwards from Infected to Susceptible, in addition to the standard transitions. We used a simple Infected-Recovered (IR) to model the behavior of the mobile marketplaces. The IR model assumes that the marketplaces start with some malware.

Mobile devices are connected to two networks: their personal connections and the marketplaces. All devices are connected to the largest marketplace, and are connected to other marketplaces in a probabilistic manner based on the relative size of the marketplace to the largest marketplace. Connections between nodes are represented by weighted and directed arcs. The weight for incoming arcs, is equal to the number of communications received by a given neighbor divided by the out degree of that neighbor. This represents the likelihood that a node will receive a communication from a given neighbor, dependent on how many other nodes the neighbor is communicating with. [140] Response to that communication is weighted by the minimum of communications received and communications sent between the neighbors, divided by the maximum of communications received and communications sent. [149]

We also assume that users, upon discovering a malicious application, will alert their neighbors to the infection, creating an inverse infection with a successful contact rate of ν_2. However, we also assume that the success of the inverse infection, as in the case of the standard infection, is weighted by the amount of contact shared between nodes.

We only model the exposure of the mobile marketplace to a single piece of malware. Based on the implementation of the malware, all connections stored in a given mobile device are revealed to the malware, representing a loss of privacy – but not necessarily the infection – of those contacts. We assume the malware is removed from a given marketplace upon the malware's discovery.

Moreover, we assume that if the malware is discovered in the official marketplace, copies of that malware are removed from all mobile devices, though reinfection can occur if other marketplaces still host the malware. This ignores the ability for users to install apps directly from other websites, but simplifies the model.

We assume a static population, with static connections to marketplaces determined at the beginning of each simulation run. The static population and static connections only approximates social pressure to change connections and the willingness of users to circumvent security measures to accomplish their own

goals through the contact rate of the third-party marketplaces.

Model Structure and Dynamics

The structure of infection for mobile devices follows the same basic structure of a susceptible $\rightarrow$ infectious $\rightarrow$ recovered model (Figure 4.2). However, there are some complications in our model.

First, it is possible for users to uninstall an app without realizing it is infectious. They may later reinstall the app due to popularity among contacts or spontaneously. This creates a loop between susceptible and infectious individuals.

Second, transition to the removed population is due to a virulence factor. We use the biological definition of virulence: removal from an infectious population due to the death of the host. [9]. We use this definition of virulence because the more aggressive a given piece of malware is, in terms of resource use, the more likely it is to be discovered and removed from a given device. Virulence also serves to remove the infection from the marketplace.

As a device discovers the malware and recovers, it reports the malicious app to the marketplace. Each marketplace responds at its given response rate to remove the malware. In the special case that the recovering marketplace happens to be the official marketplace, it attempts to remove the malicious app from connected devices. However, devices may retain the malicious app if they are connected to 3rd party marketplaces that continue to host the malware.

Finally, our model also tracks account exfiltration due to contact with infected devices. If a device i is connected to an infectious neighbor j via an arc $j \rightarrow i$, we assume that j has leaked certain contact details, such as email addresses and phone numbers, about i. If any neighbor j such that $w_{j \rightarrow i} > 0$ becomes infectious, i's information is considered exposed. Thus, privacy states follow the following transitions: Protected $\rightarrow$ Exposed and are dependent on the infection dynamics.

Model Parameters

We adapted the application discovery and adoption behaviors from different human subjects studies to parameterize our model for different populations. [3, 44, 161] Application discovery and adoption were folded into our β parameters, and differed based on user population. The social network scales β_1 by the weight of each connection $w_{j \rightarrow i}$.

We examined two different types of user behavior for smartphone app adoption:

Table 4.1: The difference in non-spontaneous infectious and recovery parameter values for Android and Apple user behaviors. Android users rely more on their social network for discovery and installing applications β_1, ν_2, while Apple users rely slightly more on download counts on the app store β_2, ν_3.

User Behavior	β_1	β_2	ν_2	ν_3
Android	0.85	0.15	0.85	0.15
Apple	0.44	0.56	0.44	0.56

- Primary reliance on social network, secondary reliance on global popularity in choosing to install the app. This type of behavior is typical of Android users.
- Roughly equal reliance on social network and global popularity in choosing to install the app. This type of behavior is typical of iPhone users.

In both of these behaviors, users have the option of removing the app without realizing that it is malicious. The β parameters are codified using human subjects research, and Table 4.1 shows the different β and ν parameter values for both Android and Apple users. We set the parameters governing the removal of an application due to malware detection (ν_{1--3}) to match the strategies for application adoption. Social network weights are held constant between groups to study effects of application adoption of the user groups, given similar disease characteristics. However, each individual in the simulation has their social response parameters scaled by the contacts they most communicate with. We also keep track of each individual's state, rather than use mean field approaches.

To compare our model with a more standard susceptible $\rightarrow$ infectious $\rightarrow$ recovered (SIR) model, we altered the μ parameter. In the normal behavior, users can uninstall an application without realizing the application is malicious. In our model of normal behavior, $\mu = \frac{5}{N_1}$, where $N_1 = 8350$ is the total number of devices in the simulation. In the SIR case, $\mu = 0$. Thus, users are 5 times as likely to uninstall a malicious app and miss the malicious behavior than to discover and report it.

Each device is given a probability of connecting to a given market, c_i. Each market is also given a probability of reacting to a malicious app report from a device, τ_i. All parameters that were used solely on devices were converted to a probability distribution at computation time.

Table 4.2: Definitions of included symbols

Notation	Definition
S_{md}	Susceptible mobile devices
I_{md}	Infectious mobile devices
R_{md}	Removed mobile devices
I_{mp}	Infected marketplaces
R_{mp}	Removed marketplaces
β_1	Effective app adoption contact rate via social network popularity
β_2	Effective app adoption contact rate via global popularity
β_3	Likelihood of spontaneously adopting app
μ	Likelihood of app removal w/o malware detection
ν_1	Likelihood of spontaneous app removal with malware detection
ν_2	Likelihood of social contact app removal with malware detection
ν_3	Likelihood of app removal due to market response
$w_{j\to i}$	Weight of connection from subject j going to subject i.
c_i	Probability of a connection to market $i \in N_2$
τ_i	Probability of reacting to a malicious app report. i
MP_O	Official marketplace
N_1	$S_{md} + I_{md} + R_{md}$
N_2	$I_{mp} + R_{mp}$

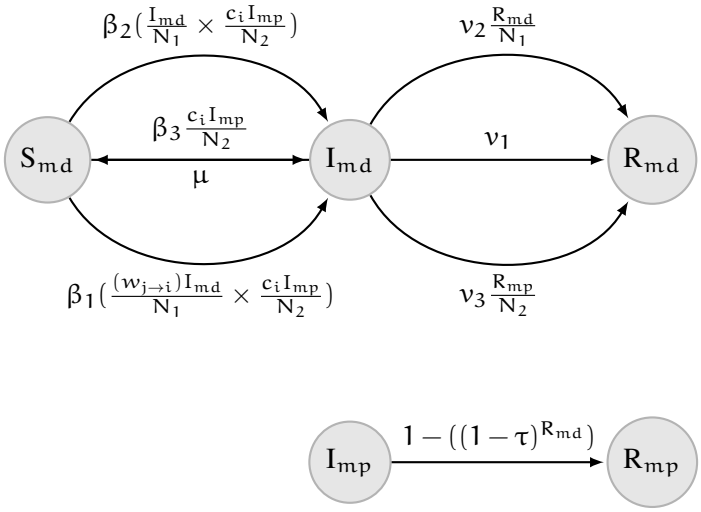


Figure 4.2: Allowable transitions in our two-population SIS model with recovery and social response

Model Analysis

We ran our simulations 1000 times over four levels of user behavior (Android and Apple users; and Android and Apple with $\mu = 0$) and five different market place configurations (Apple Market, Small Android Market, Android Market, Nonvendor Market, and a market that behaved as a standard SIR model). We then used a hierarchical Bayesian model to explore two infection quantities and one privacy quantity on the different user behavior and market place configurations. We examined a growth rate parameter based on the total number of infected devices at the end of the simulation run (R0), the total duration of the infection, and a growth rate parameter based on the the total accounts exposed at the end of the simulation run (P0).

4.4 Results

We verified the functionality of our Android malware with a testbed server and with human subjects. We found that with the human subjects used the program long enough to avoid the bandwidth and runtime limitations. Due to the limited number of participants, we had to adjust the smart phone contact network that we used for our modeling simulations.

We had two key results from our simulations. While user behavior was a factor in limiting the spread of malicious apps on official markets, a well functioning official market place was able to effectively prevent the spread of malware. User behavior, however, was a major component of limiting malware spread on markets dominanted by non-vendors. While a well functioning market place was able to mitigate the spread of the malware, malware of the same type that we developed, was able to leak contact information for the entire giant component faster than the market place or users could respond.

Experimental Results

We installed and tested our app on an Android 4.0 device which contained over 200 contacts to test the functionality, speed, and data transfer of our application. We found that, on average, approximately 1 kilobyte is required to send each contact's information across the wire along with some minimal formatting to make the transmission easily readable in the server's console. This could be reduced rather easily by simply removing the formatting that is performed on the phone and move it to the server.

Android requires that any Internet communications conducted during the operation of an app are performed in an execution thread separate from the main functionality of the app. This requirement actually

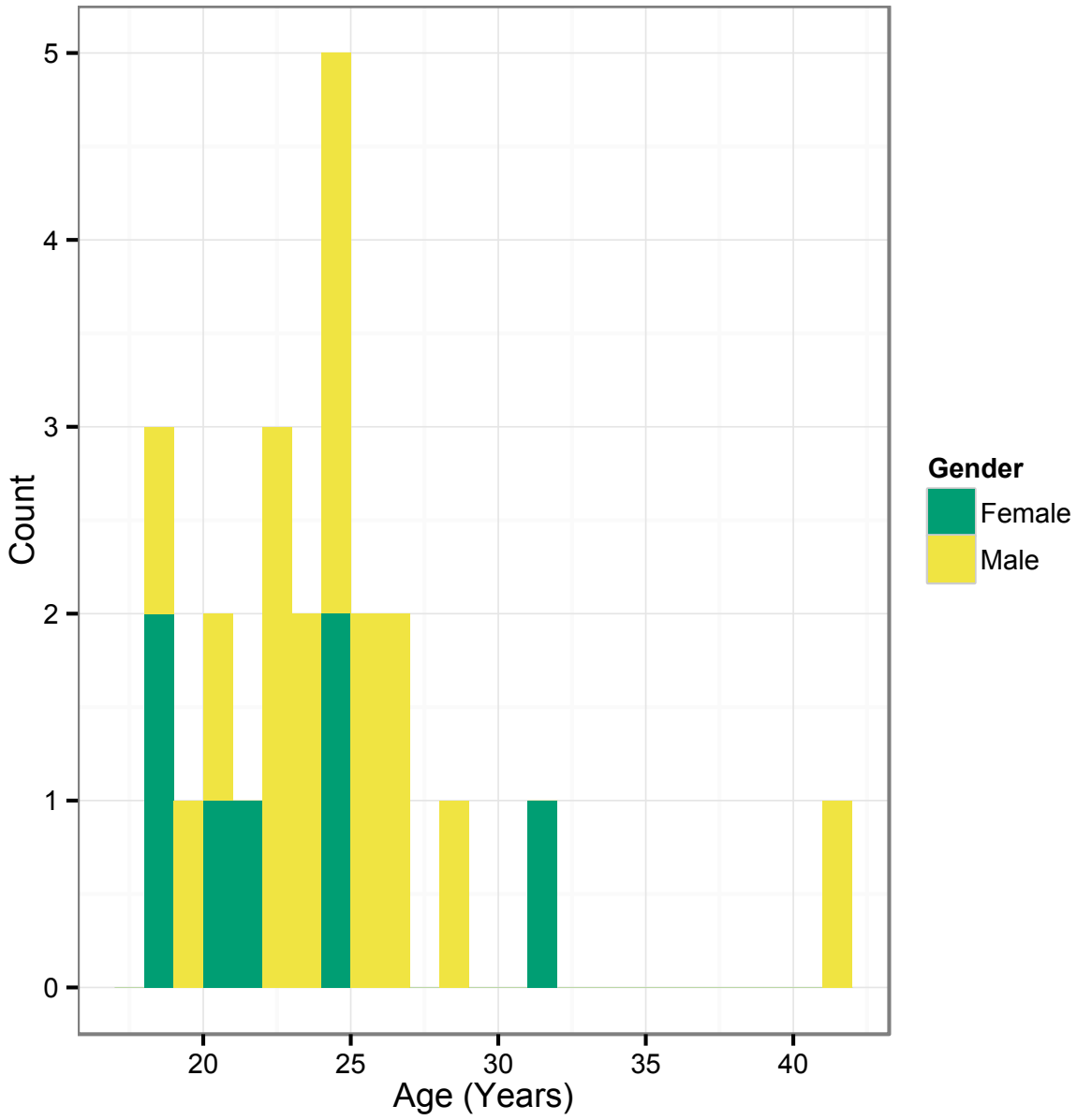


Figure 4.3: Age and Gender Distributions of Participants that Elected to Install our Score-Keeping App

helped us to keep malicious activities operating in the background, unknown to the app's user. The separate execution threat prevents the main, legitimate functionality of the app from being affected by the execution of the malicious code.

After the test of the code on our testbed server, we surveyed 116 participants to test the application. Of the 116 participants we asked to install our application, only 24 subjects agreed to install the application (Figure 4.3). We collected an additional 4106 unique contacts from the 24 participants (Figure 4.6).

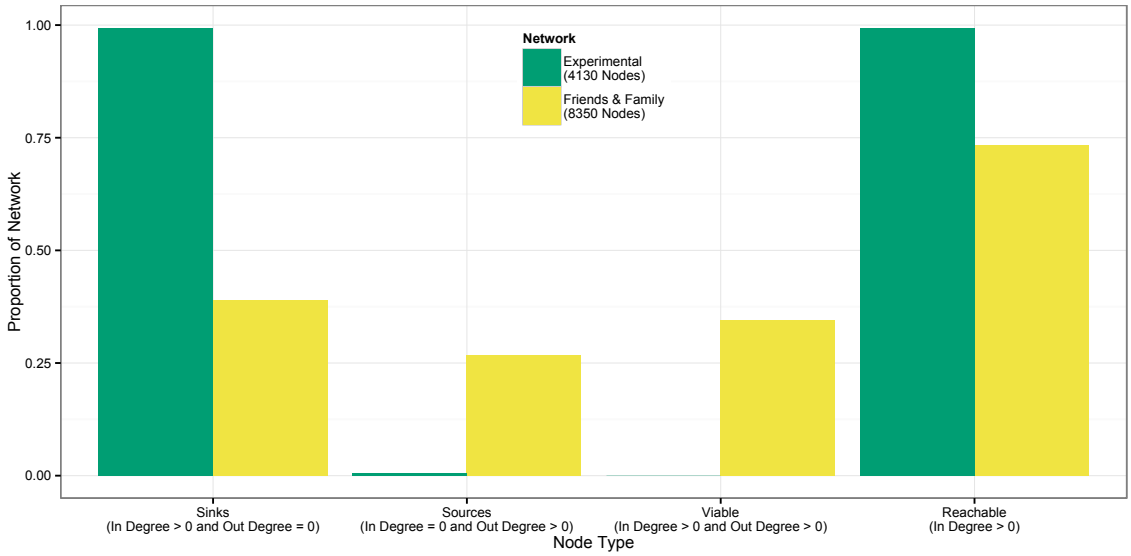


Figure 4.4: Comparison between the node type in our experimental network and the Friends and Family network.

While there are numerous subjects with very few contacts, and pose little risk to the network at large, those subjects with large numbers of contacts threaten to expose an significant portion of the network's contact information, if they are infected. For example, the subject with the maximum number of contacts, if infected with our malware, would leak contact information for 921 of the 4106 contacts in the network, as well as acting as a super-spreader of the malicious application.

Given, however, that we had only 24 subjects, and only retrieved contact information from their address books, our information of the contact network is quite limited (Figure 4.5). For example, there were no viable nodes (out-degree > 0 and in-degree > 0), only sources and sinks (Figure 4.4). We do not assume that a subject's contacts also had that subject in their address book as well. Furthermore, we have limited information about the structure due to the fact that we did not collect information from a subject's contact's contacts, nor do we have information about untested entities that have the subjects as contacts.

Simulation Results

The limitations of our experimental protocol required us to use a more robust network for looking at mobile application adoption. Using this larger, more-connected network, different user behaviors, and market place configurations, we found that user behaviors were very modest in controlling malware spread, but a well-functioning central market place, even with active third-party markets, is effective at mitigating large-scale

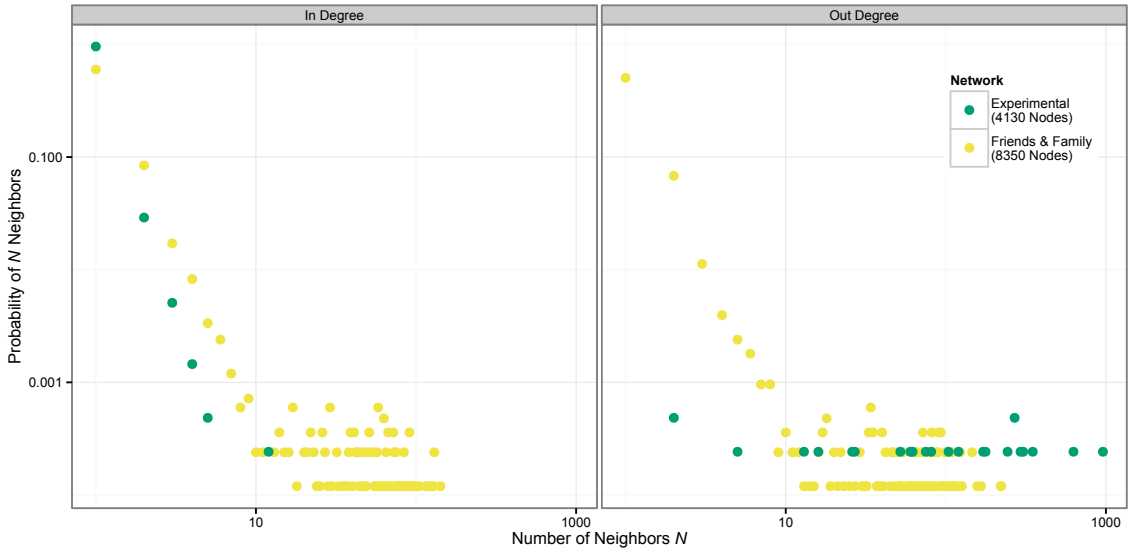


Figure 4.5: Comparison between the degrees in our experimental network and the Friends and Family network.

spread of malware. However, neither user behavior, nor market place configuration was able to affect the loss of privacy information from mobile malware.

Network Analysis

Due to the limitations in our experimental social network (Figure 4.6), we used the Friends and Family (FF) dataset, a social network that had been experimentally determined over three months with 185 participants (Figure 4.7). The network includes number of calls and SMS messages between participants, as well as between participants and entities outside of the participants' network [3].

Figure 4.4 shows some of the numerical differences in number of nodes and edges. In particular, the FF network has a more uniform distribution in terms of node types, while our experimental network is, aside from the initial subjects, sinks. The out-degree distribution comparison in Figure 4.5 further quantifies the differences in network structure. In our experimental network, while the majority of subjects have low numbers of contacts, one-third of the subjects have more than 200 contacts. In the friends and family network, only two of the nodes have more than 200 contacts.

The network our experiment generated was based on address books, while the FF dataset was based on actual communication between entities. This difference partially explains the variance in network structure, but there are also demographic differences between the networks. The Friends and Family study specifically

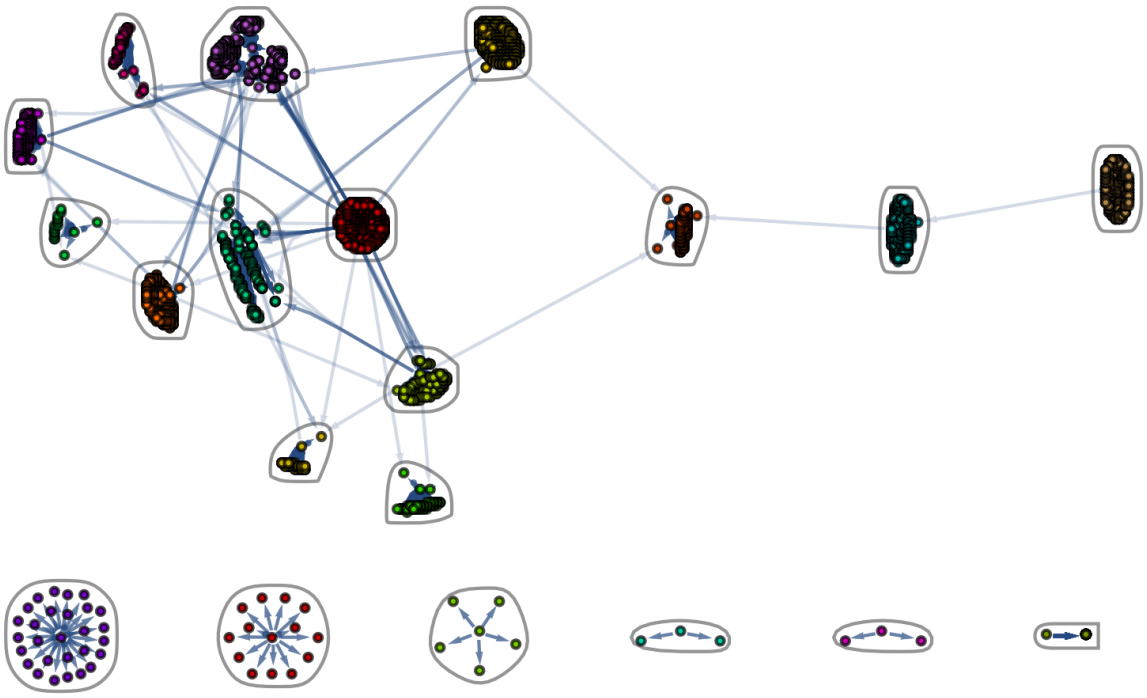


Figure 4.6: Graphical representation of the community structure in the social network extracted from our experimental data. Number of nodes = 4130.

targeted young families, while our work was conducted on a university campus. [3]

The network analysis of the FF network demonstrates how many entities a subject realistically maintains contact with, rather than how many contacts are stored in their address book. The differences are not contradictory, however, as other research shows that the strength between ties is smaller when one has fewer contacts, and strong ties are related to application adoption and social phishing. [3, 106, 140] Thus, the FF network works well for our modeling because, while it does not capture the entirety of a subject's address book, it does portray the connections a subject is most likely to respond to in regards to installing a mobile application.

Simulation Analysis

We used the FF network to simulate a group of smart phone users exposed to a malicious mobile application. We looked at the effect of user behavior and market place structure on the spread of the malicious mobile app. We found that market places were effective mitigators of malware spread. User behavior had little effect in well functioning market places, but was critical to limiting malware spread when the market place

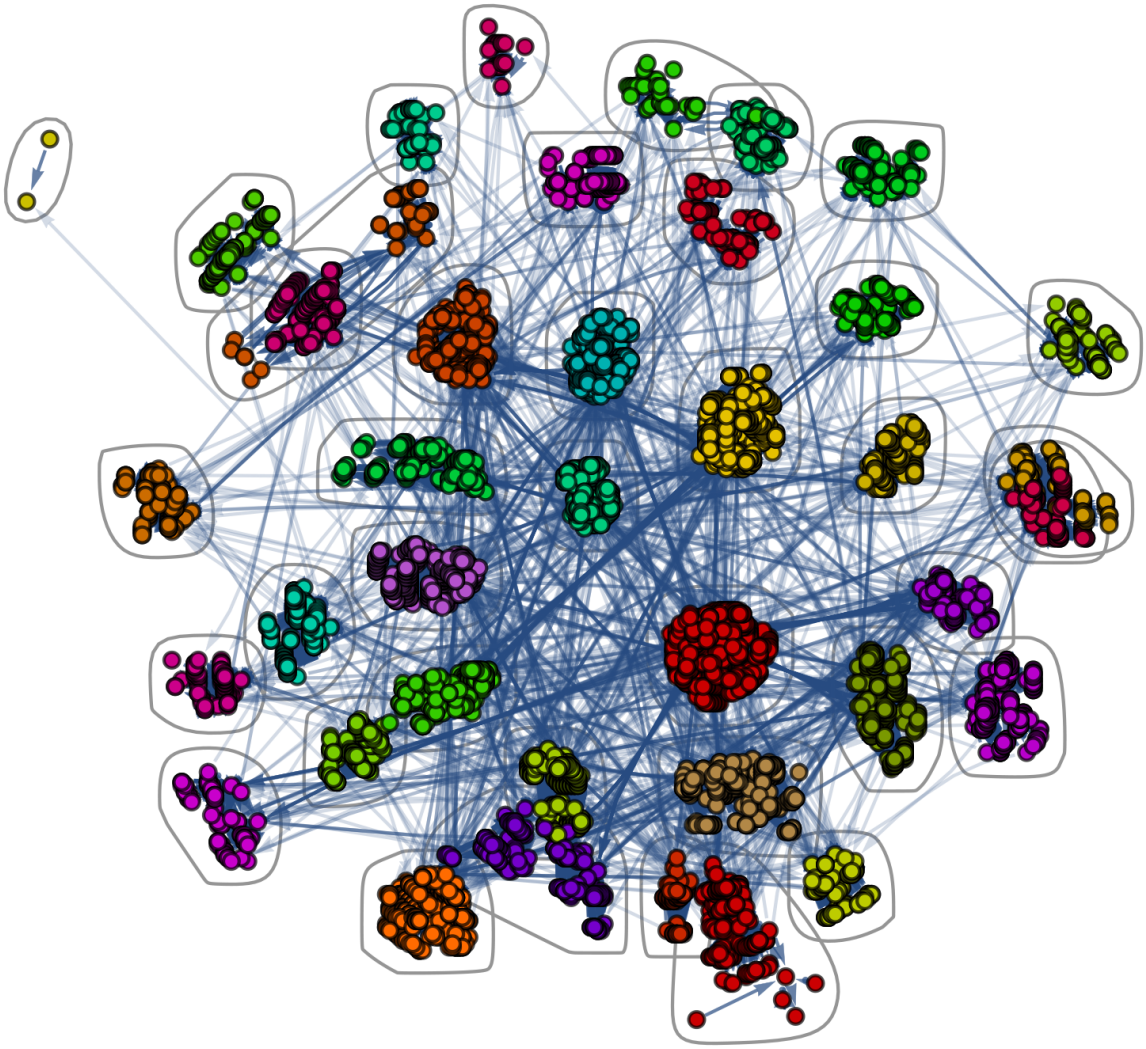
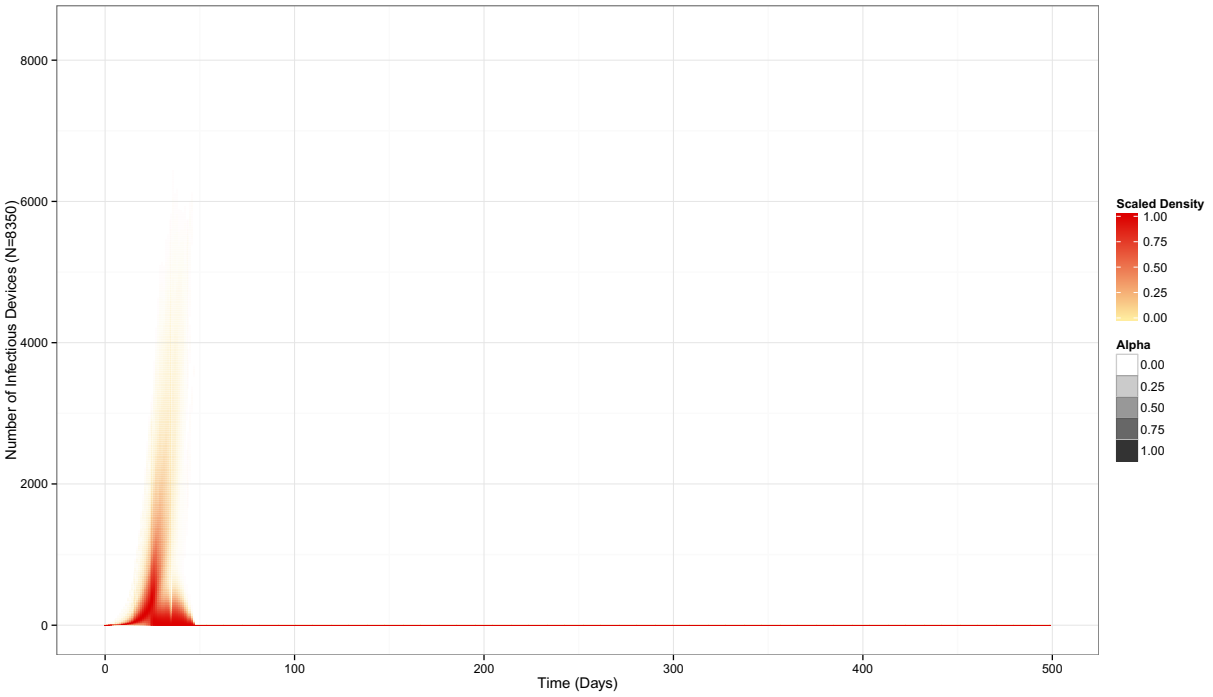


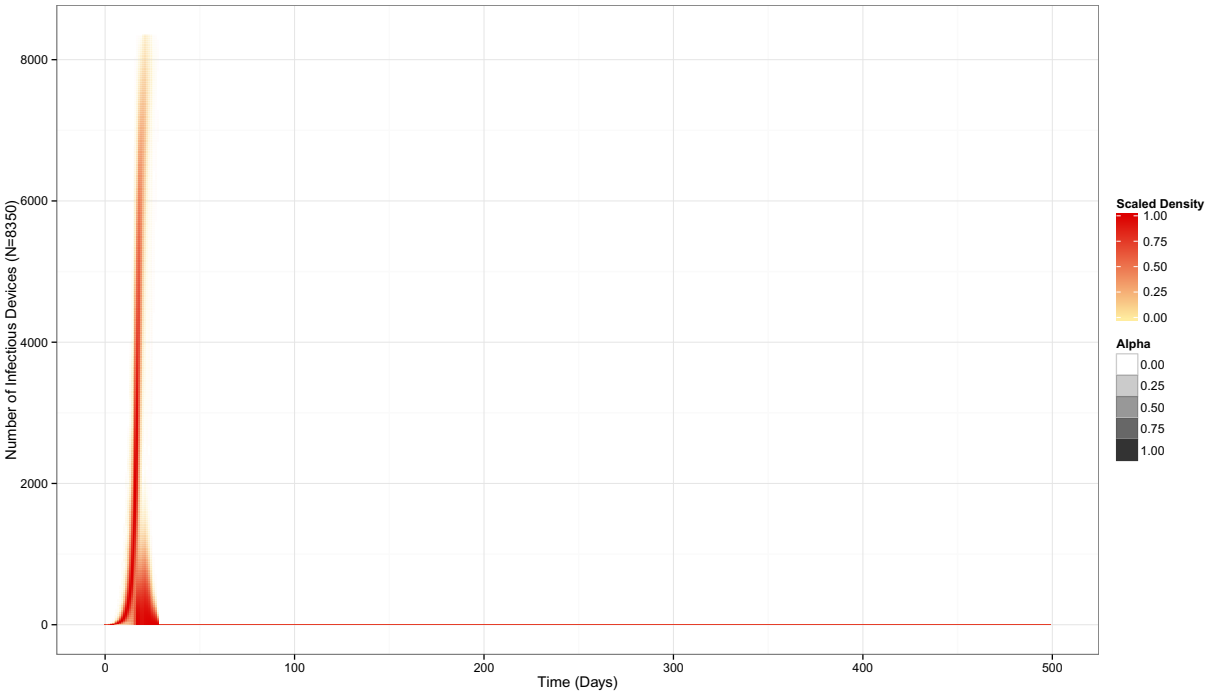
Figure 4.7: Graphical representation of the community structure in the social network extracted from the Friends and Family data set. Number of nodes = 8350.

was not functioning as designed. Exfiltration of a agent's contacts was also mitigated by user behavior, but the effect of that mitigation is less effective as market places become less controlled.

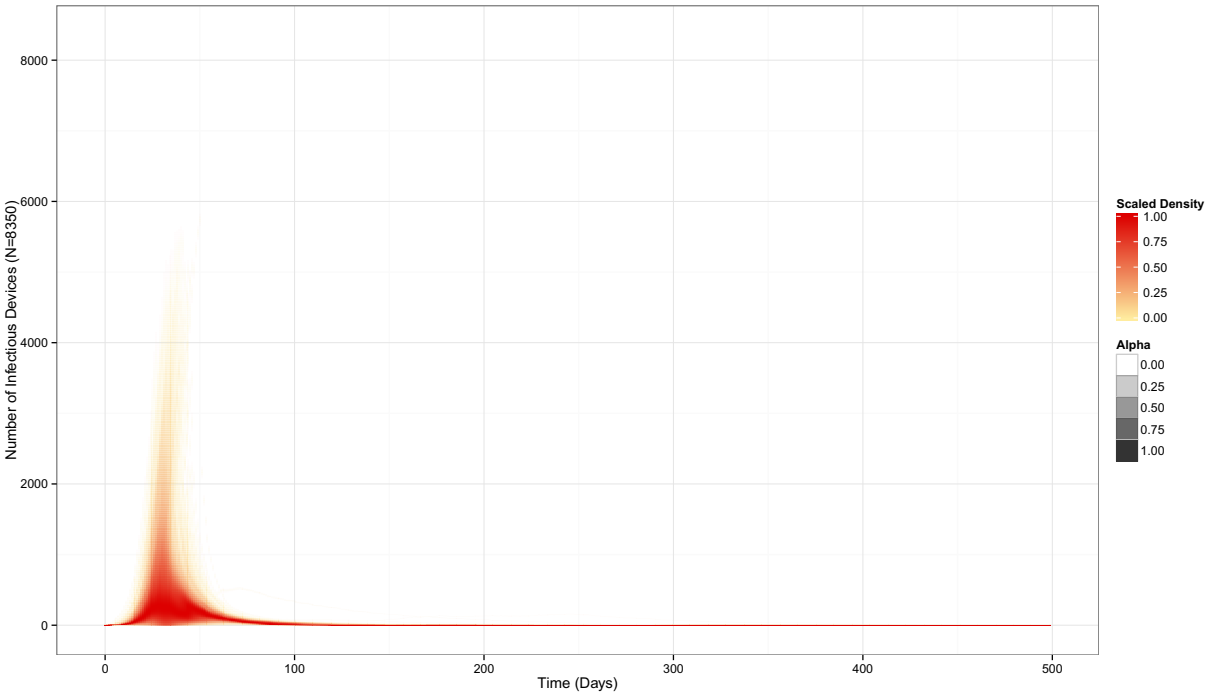
A quick look at the infection dynamics (Figure 4.8) reveals the general idea of the effects of user behavior and market structure. Across each market structure, behavior that relies primarily on one's social network exhibits fewer peak infections. After the jump from a single market, there appears to be little change in the infection duration, except when moving from a loosely regulated market (Nonvendor) to a completely unregulated market (SIR) in the Android condition, where the force of infection and recovery jump sharply in comparison to other Android conditions.



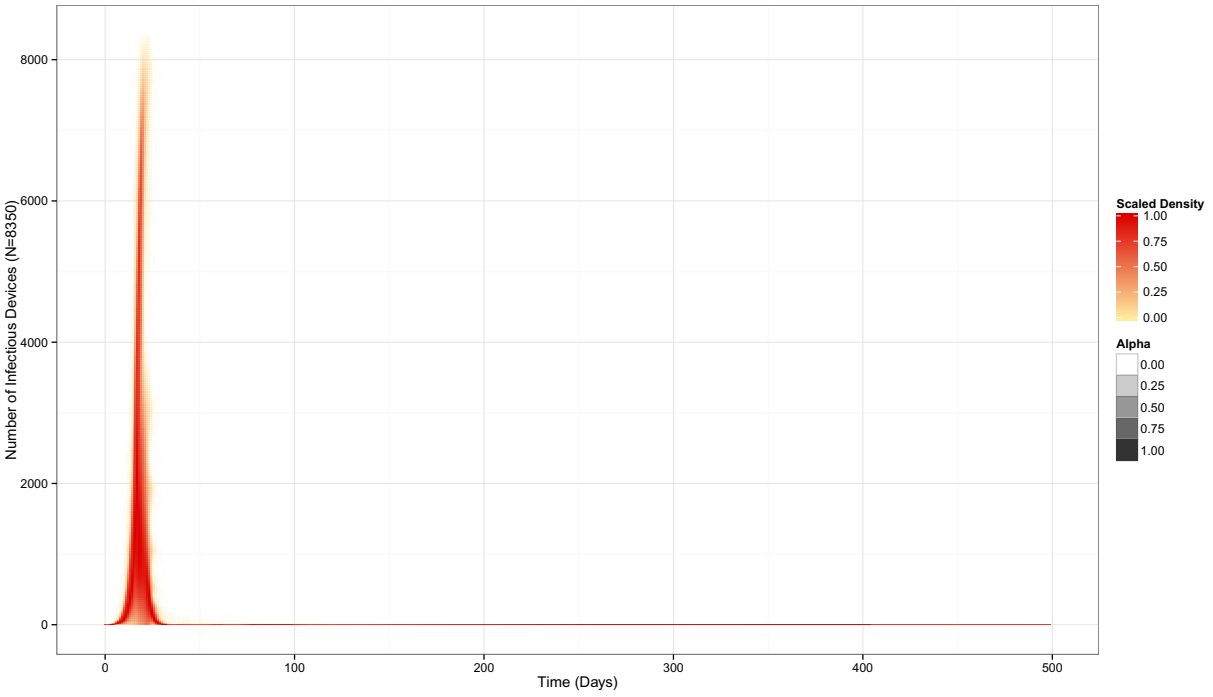
(a) Density plot of number of infectious individuals with a preference for social recommendations on a single, centralized market over 1000 simulations



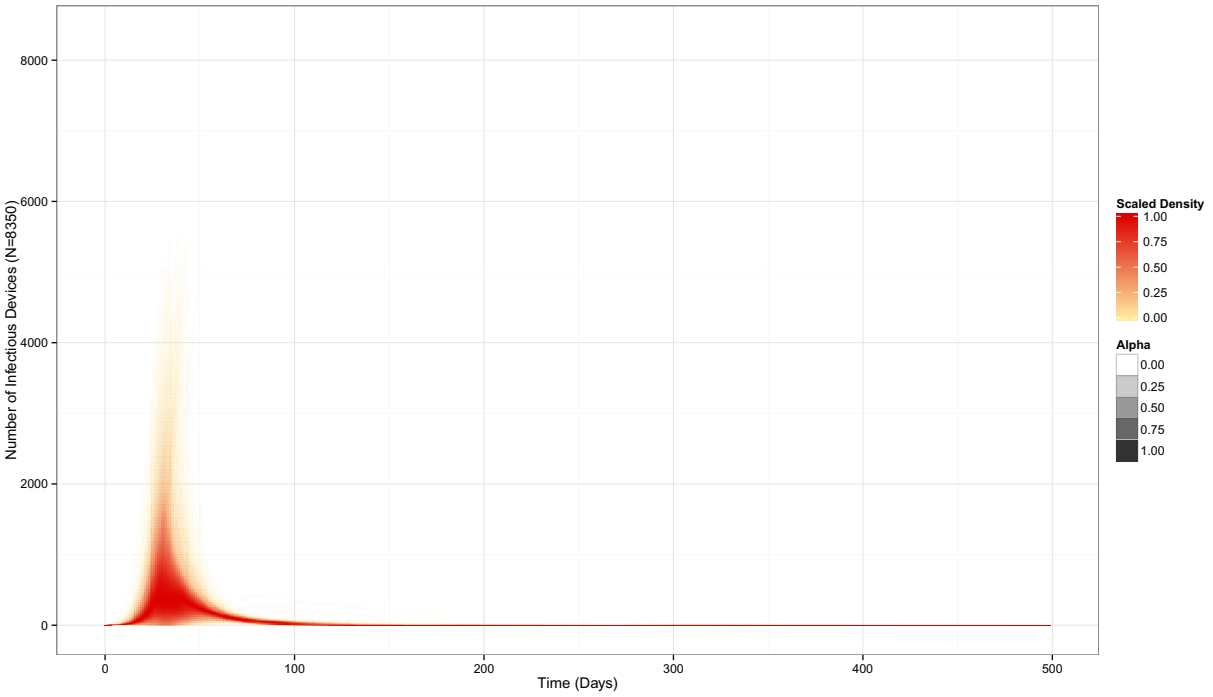
(b) Density plot of number of infectious individuals with a preference for global download counts on a single, centralized market over 1000 simulations



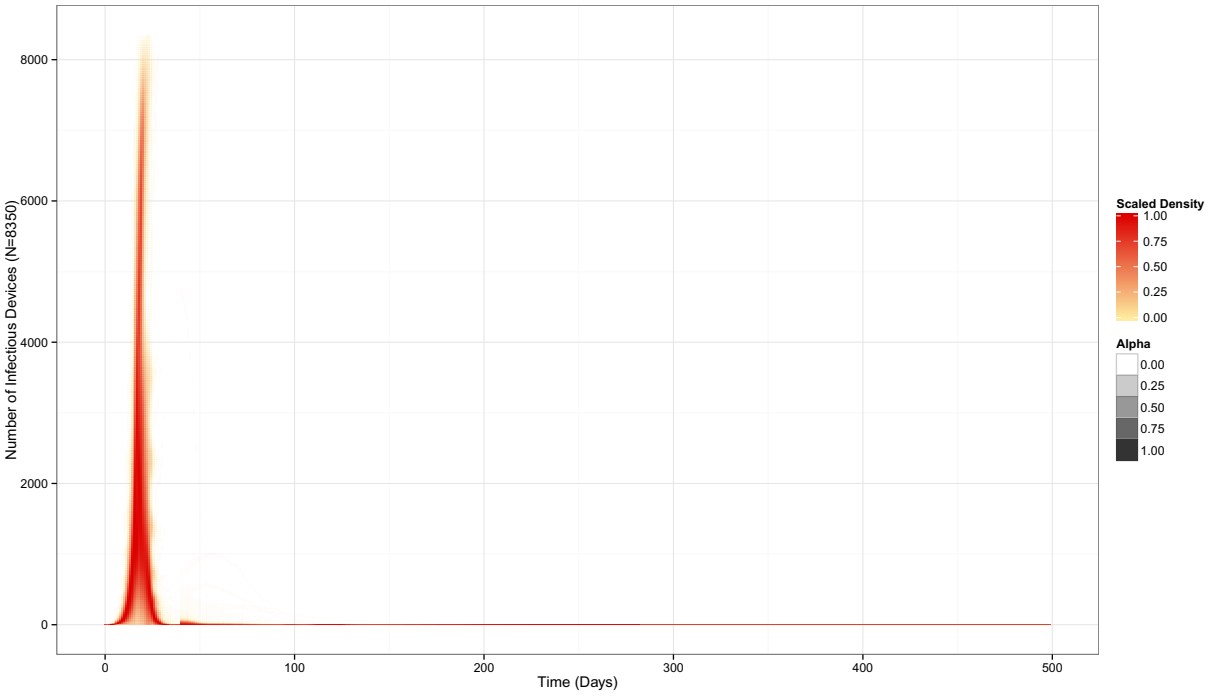
(c) Density plot of number of infectious individuals with a preference for social recommendation with one centralized market and 7 third-party markets over 1000 simulations



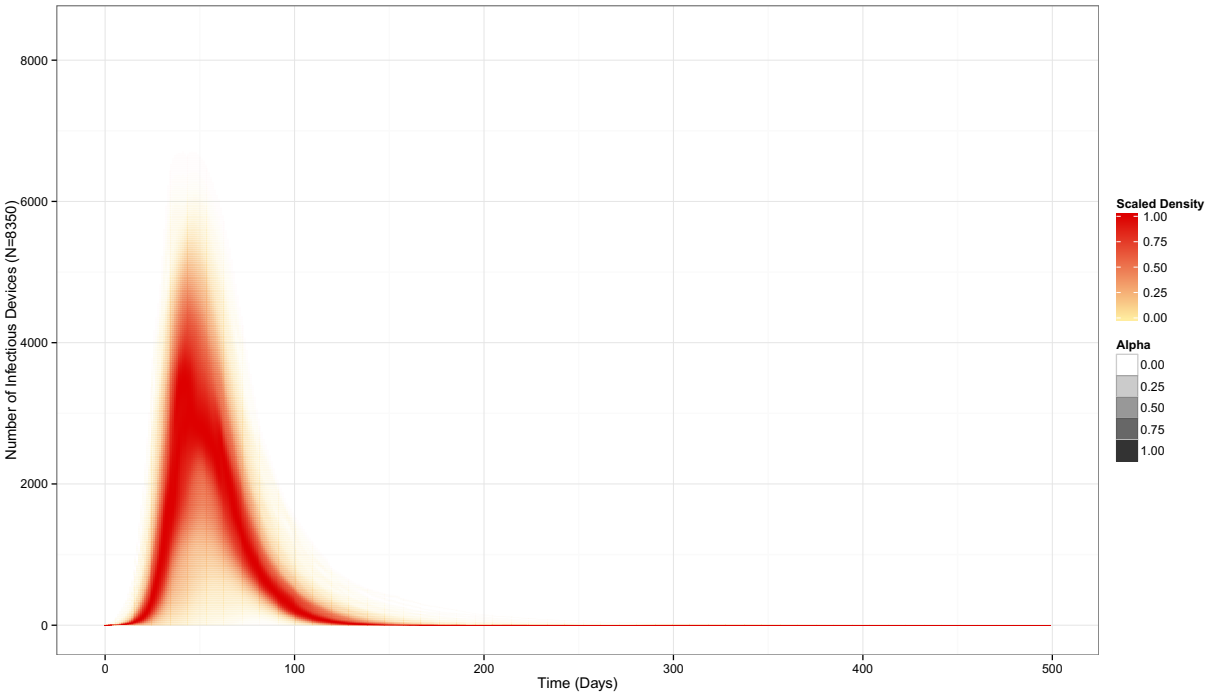
(d) Density plot of number of infectious individuals with a preference for global download counts with one centralized market and 7 third-party markets over 1000 simulations



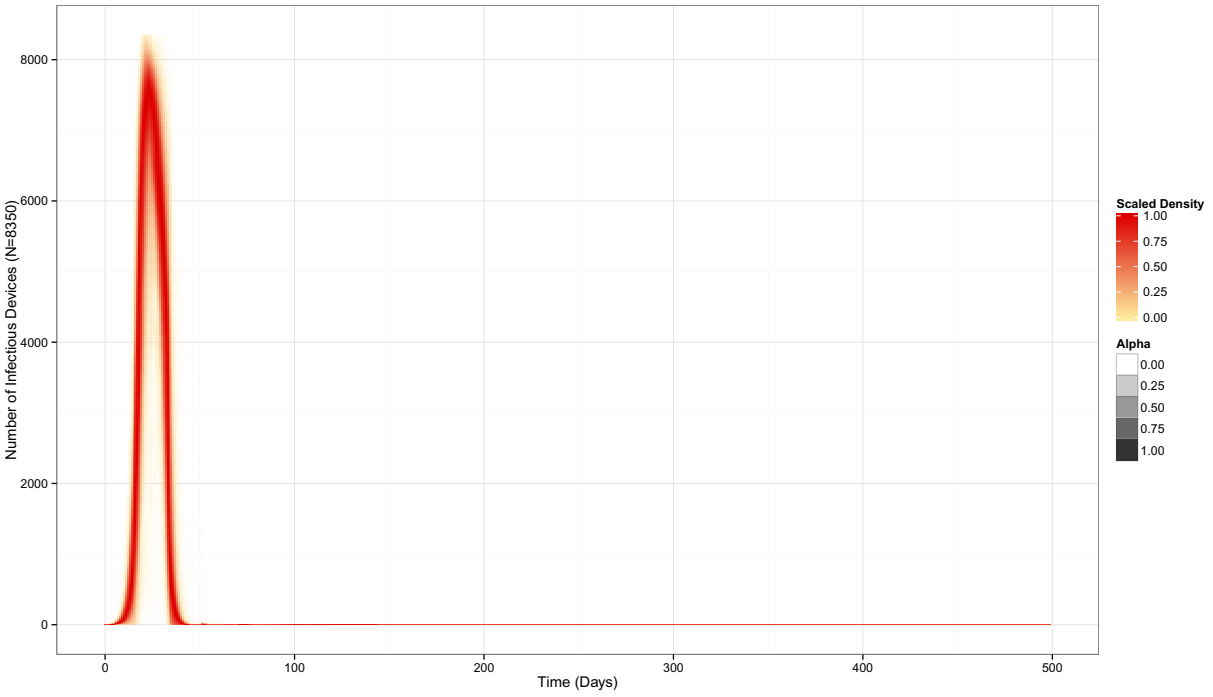
(e) Density plot of number of infectious individuals with a preference for social recommendation with one centralized market and 31 third-party markets over 1000 simulations



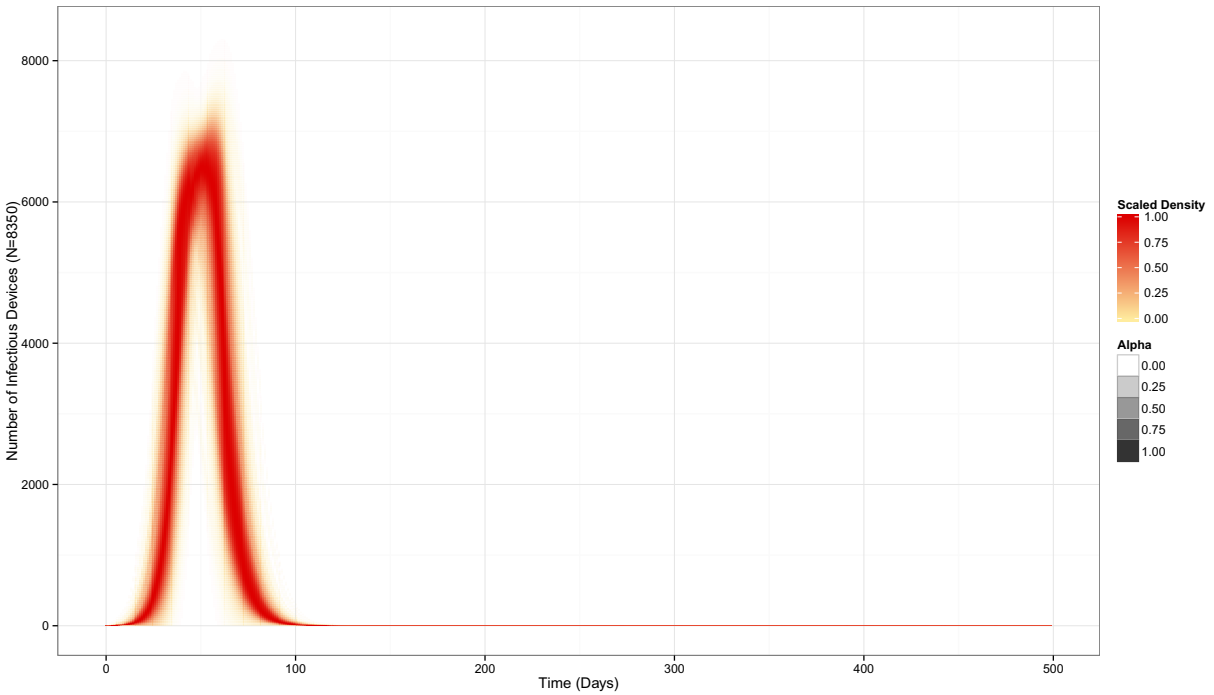
(f) Density plot of number of infectious individuals with a preference for global download counts with one centralized market and 31 third-party markets over 1000 simulations



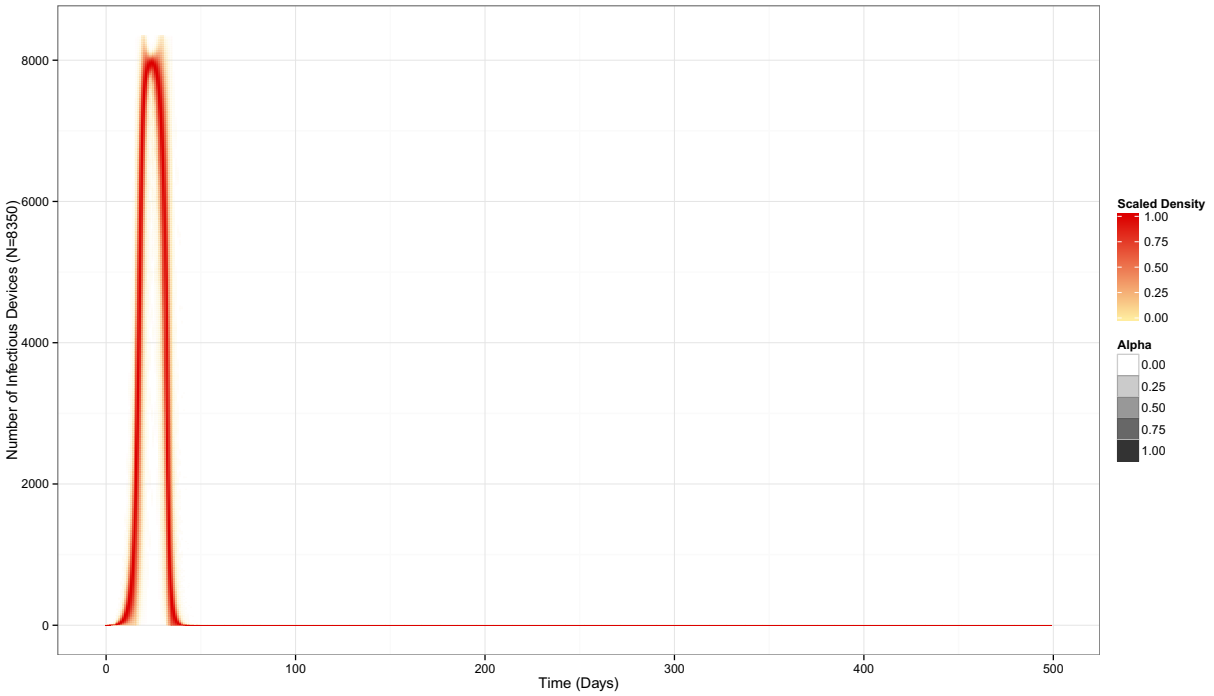
(g) Density plot of number of infectious individuals with a preference for social recommendation with 32 third-party markets over 1000 simulations



(h) Density plot of number of infectious individuals with a preference for global download counts with 32 third-part markets over 1000 simulations



(i) Density plot of number of infectious individuals with a preference for social recommendation with completely non-responsive markets (SIR)



(j) Density plot of number of infectious individuals with a preference for global download counts with completely non-responsive markets (SIR)

Figure 4.8: Visually Weighted Density of Infection Dynamics on 1000 Simulations/Condition. The Left Column Shows Android-Type Behavior and the Right Column Shows Apple-Type Behavior. The Market Structure Moves from Most Controlled (Apple) at the Top to Least Controlled (SIR) at the Bottom.

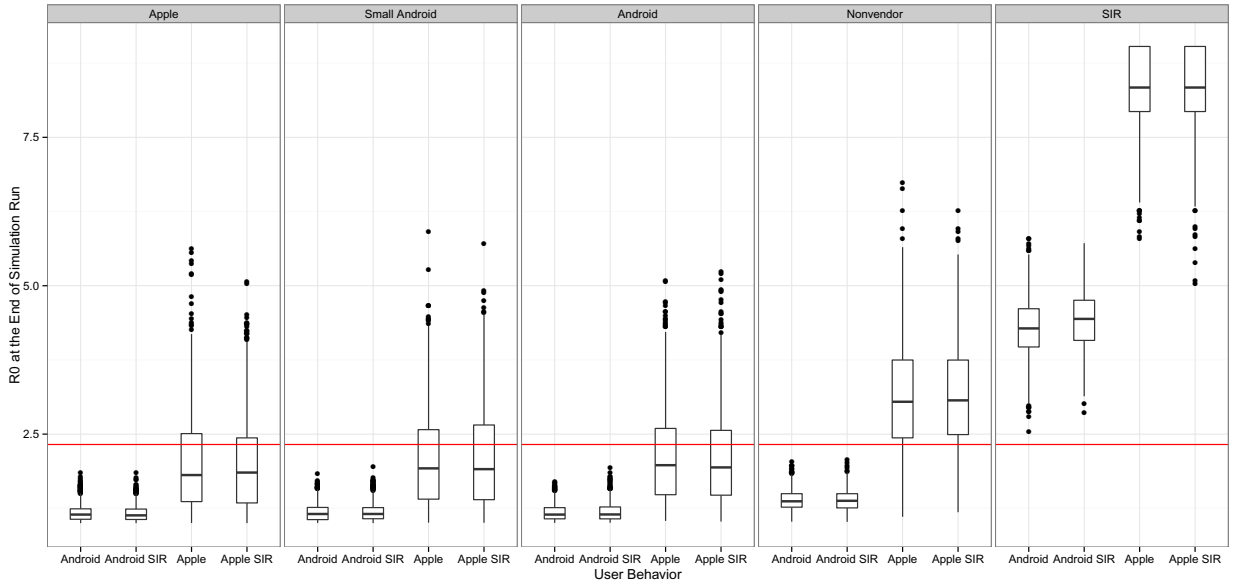


Figure 4.9: R0 at the end of Simulation Run by User Behavior and Market Structure. Markets are Ordered by Level of Centralized Control. Red Line Represents Mean of the Data.

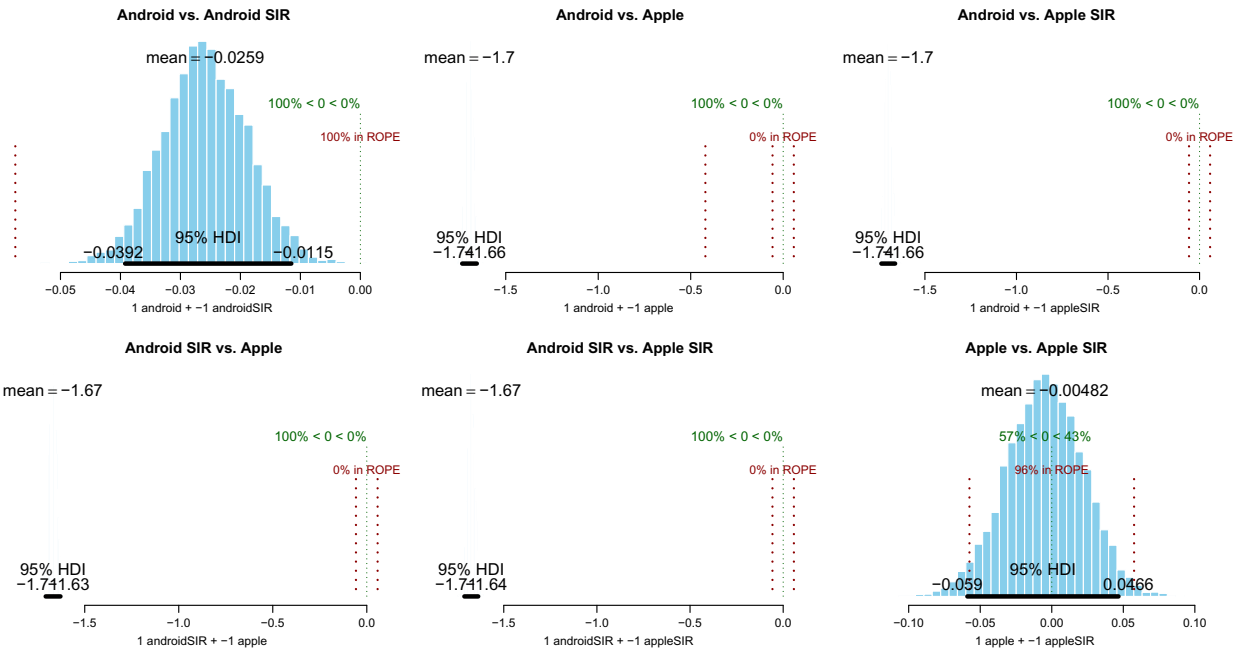
The rest of this section will examine the specific effects of user behavior and market place structure on the total number of devices infected (R0), infection duration, and total number of accounts exposed (P0). R0 and P0 are calculated by Equation 4.1

$$-\frac{\log(1-x)}{x} \tag{4.1}$$

with x = total proportion infected/revealed by a given time, t. We use the last time step of the simulation as the t in this equation. Infection duration is longest contiguous block of time where at least one device is infectious. In situations with multiple markets it is possible to have resurgent infections, but we did not calculate information on these.

R0 We use R0 as a measure of the force of infection calculated based on the number of devices that were not infected during a simulation run. Based on this measure we find that well functioning markets are effective at limited the spread of malware in comparison with a standard SIR type situation (markets take no action to remove malware) (Figure 4.9). Even nonvendor markets, those without the ability to remove malware from infectious devices, are effective at limiting the spread of malware, particularly if users are relying on their social networks to make app adoption choices.

We find, perhaps obviously, that more controlled markets are the best at limiting malware spread. However, the differences between well functioning markets are limited (Figure 4.10b). For example the differ-



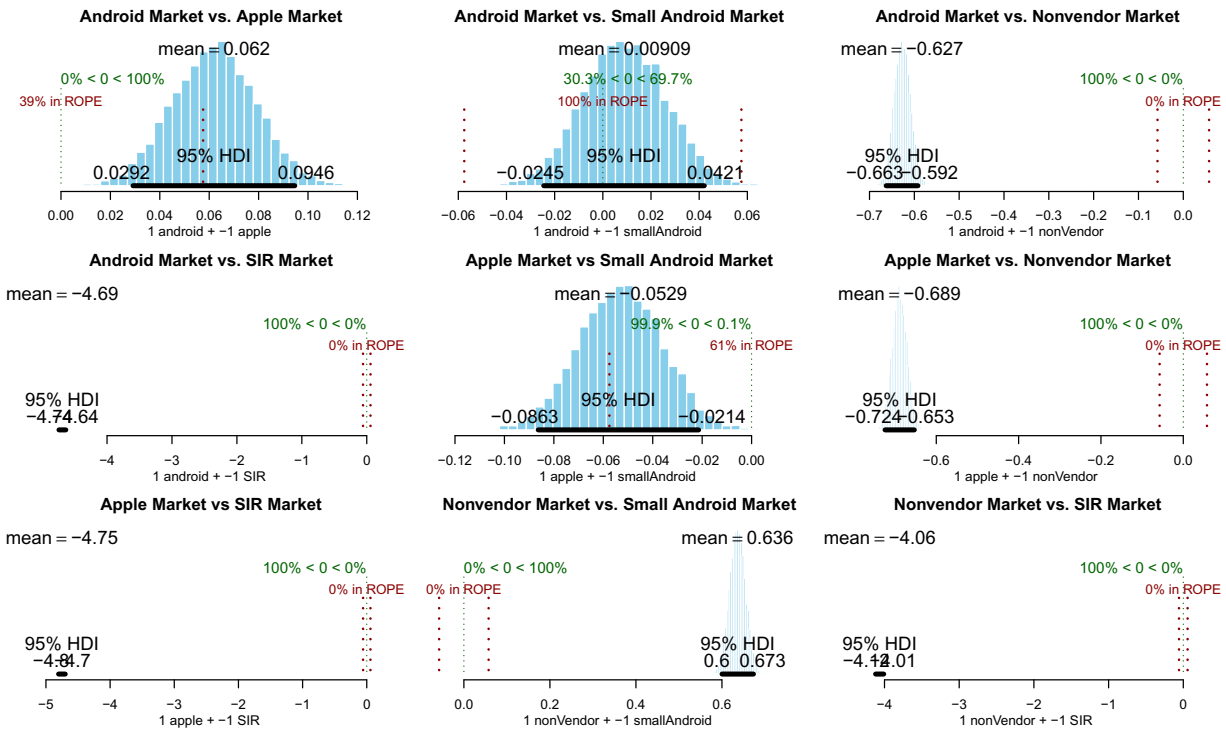
(a) Comparisons of the Effects of User Behavior on R0.

ence between the larger Android market and the Apple markets is a full order of magnitude smaller than the difference between Android user behavior and Apple user behavior (Figures 4.10a and 4.10b). The difference between the smaller Android market and the Apple market is negligible (Figure 4.10b).

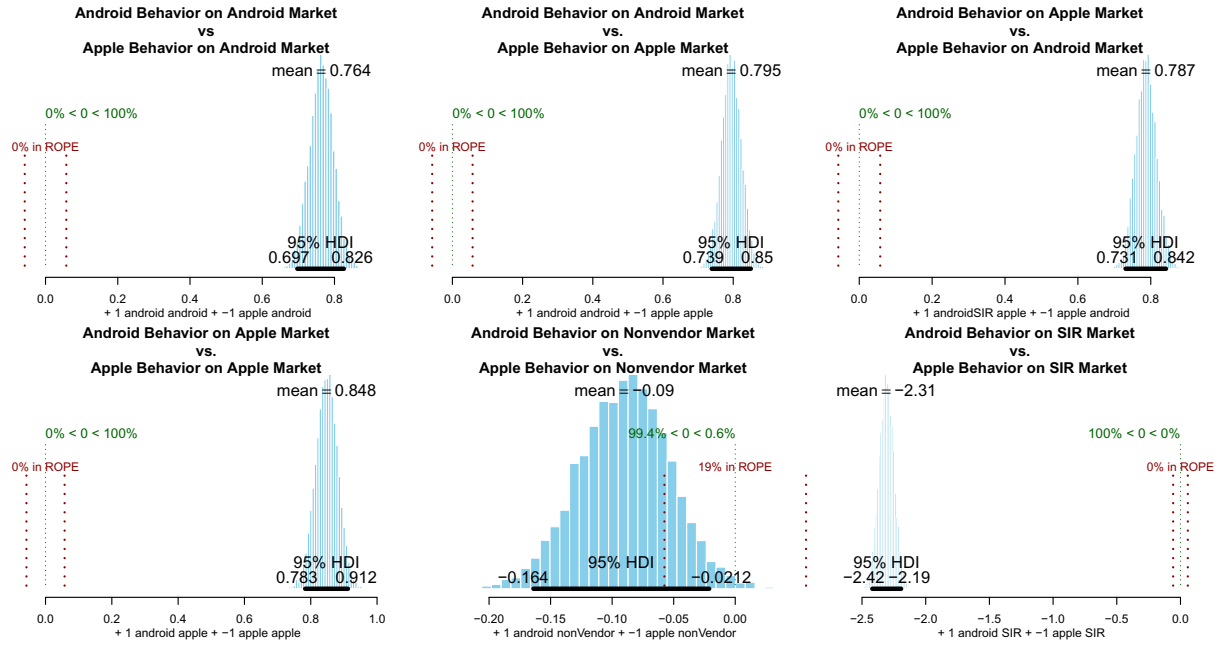
We also see the importance of market structure in the interaction effects. We find that Apple user-behavior benefits the most from the most controlled markets, and causes large jumps in infections when the ability to remove malware from infectious devices disappears and again when markets do not remove malware (Figure 4.10c). Infection duration, on the other hand is more difficult to piece together.

Infection Duration We find, unlike in the R0 and P0 cases, that both user behavior and market implementation have equal importance in affecting infection duration. There is an interesting interplay between moving between a single market to multiple market places. In general, reliance on one's social network slows the spread of malware, but also extends the infection duration as information about the infection percolates through the network (Figure 4.11).

Similarly, moving between a single market to multiple market places, allows the malware to exist in smaller markets. Thus, as the primary markets are removing malware, late adopters begin installing from less responsive 3rd-party markets. Even the simple jump from a single market to an eight market structure is significant, accounting for roughly half of the increase of duration between Apple and Small Android



(b) Comparisons of the Effects of Market Structure on R0.



(c) Comparisons of Interaction Effects on R0.

Figure 4.10: Contrasts of Behavior (4.10a), Market (4.10b), and Interaction (4.10c) Effects on R0.

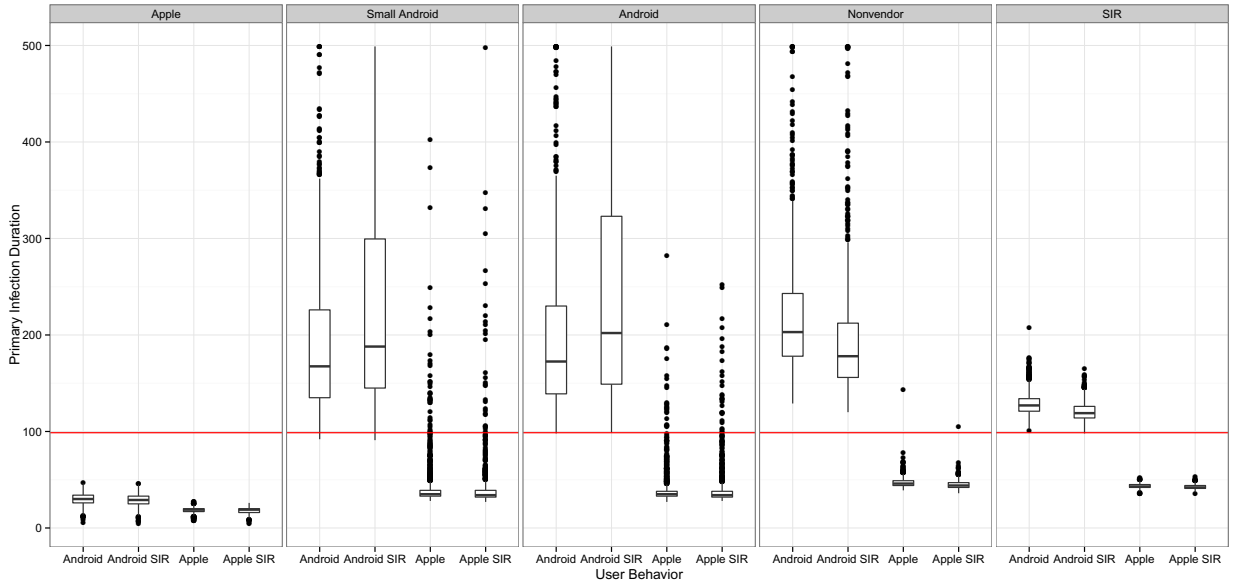


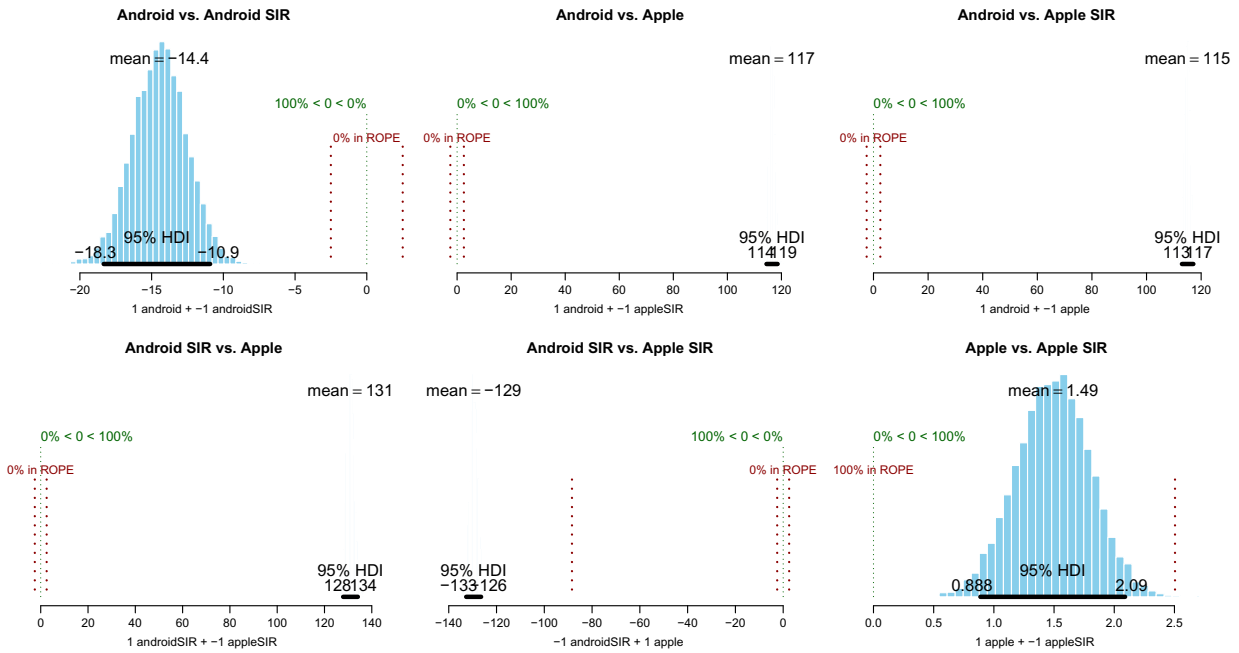
Figure 4.11: Primary Infection Duration by User Behavior and Market Structure. Markets are Ordered by Level of Centralized Control. Red Line Represents Mean of the Data.

marketplaces (Figure 4.12b). Yet there is a slight decrease in duration in the shift from a nonvendor market to the SIR market type for Android users.

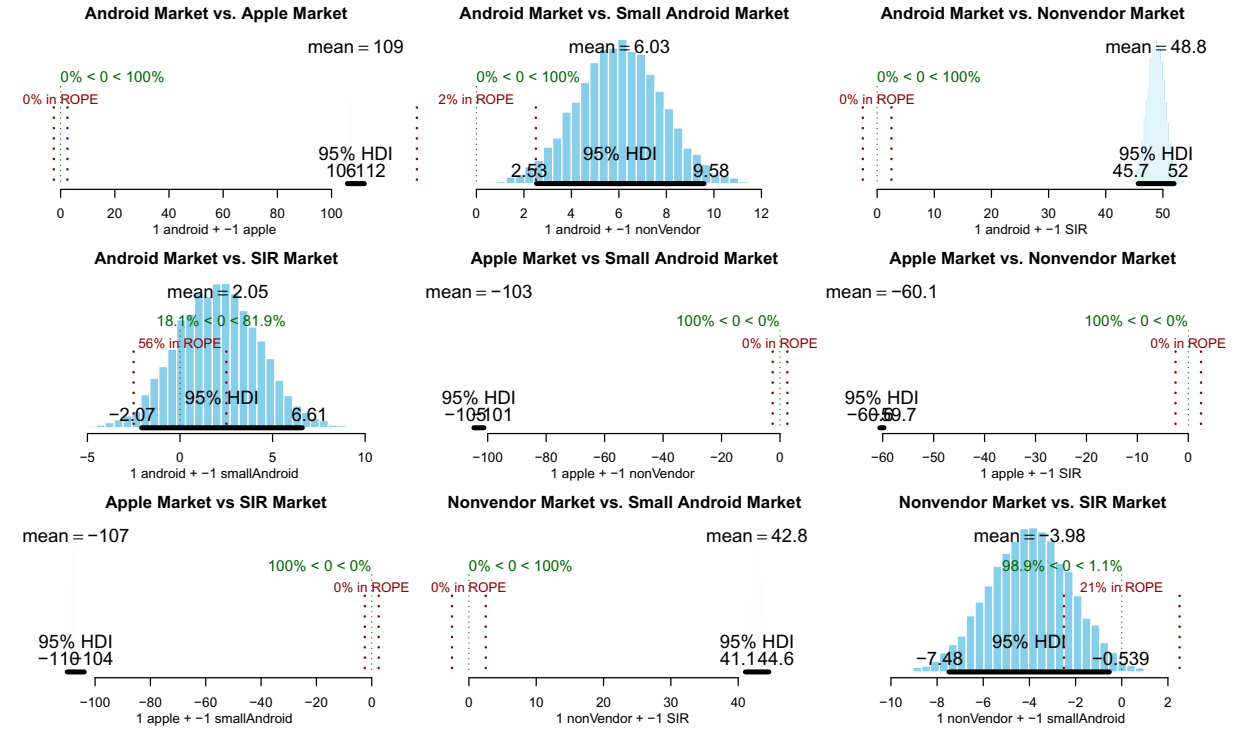
The interaction effects on infection duration are interesting. Apple behavior, in general, is better for lower infection duration. But, in most instances, Android type-behavior receives the most benefit from a given market configuration. However, the full Android market and the full nonvendor market structures further extend the infection duration of Android-type behavior.

Based on the dynamics of each condition (Figure 4.8), with the exception of the Apple market, it appears that infection duration is related to how quickly the malware can spread within the marketplace before the marketplace responds to the infection. The Apple-type behavior has shorter durations, because more of the network is compromised faster.

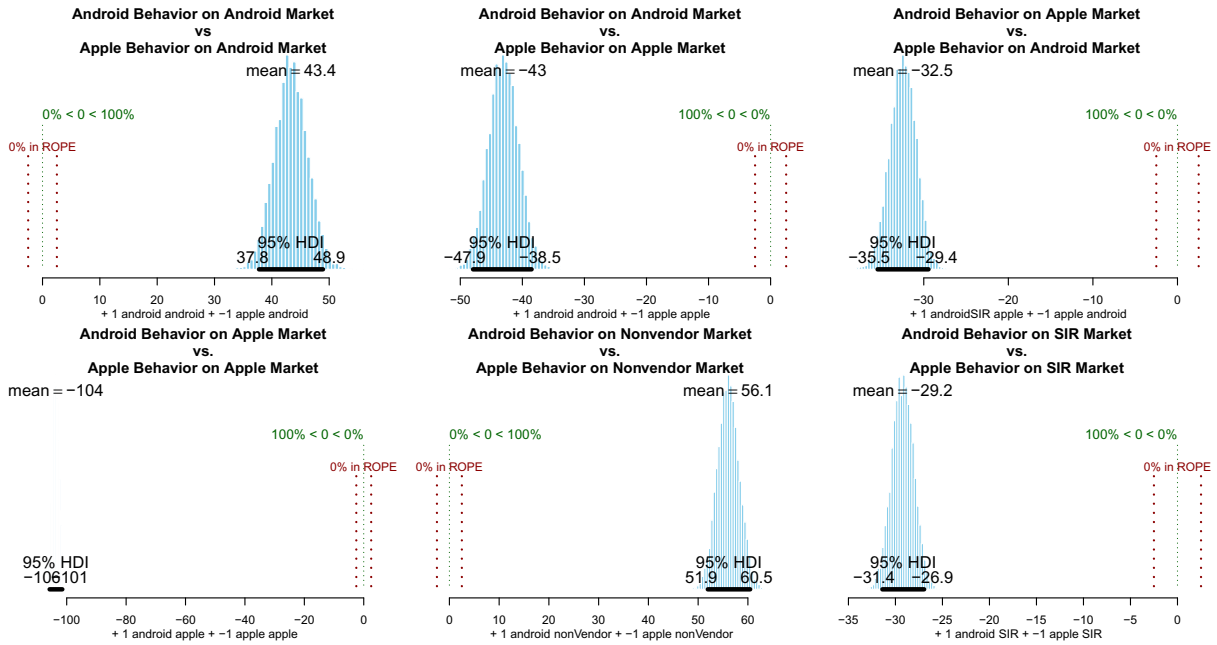
Thus, the infection duration and peak dynamics are examples of how quickly a given piece of malware spreads–an important aspect of severity–but it does not change the total number of affected individuals. Rather, it is evidence of a much more rapid spread. In this case, our model has a virulence factor, allowing for particularly forceful infections to be caught, and then quickly eradicated after affecting most, if not all the population. The slow curves of Android-user type behavior represents more of a propagation type of spread, while the sharp peaks in Apple-user type behavior are reminiscent of common source type infections (e.g., poisoning the well).



(a) Comparisons of the Effects of User Behavior on Infection Duration.



(b) Comparisons of the Effects of Market Structure on Infection Duration.



(c) Comparisons of Interaction Effects on Infection Duration.

Figure 4.12: Contrasts of Behavior (4.12a), Market (4.12b), and Interaction (4.12c) Effects on Infection Duration.

P0 The results for the number of accounts revealed follow the same general pattern as those for number of infected devices (Figure 4.13). However, the importance of user behavior was more intense (Figure 4.14a). Similarly, the market effects are also more important in reducing the number of accounts that are revealed (Figure 4.14b). The interactions, in contrast, are less important in determining account loss than those found for R0 (Figure 4.14c).

There are two exceptions, however. When we examined R0, there was very little difference between the full Android market and the small Android market (Figure 4.10b). In regards to the exfiltration of neighbor data, however, even the suggested small difference between the full and small Android markets has disappeared (Figure 4.14b). Based on the comparisons between the Apple market and the two Android markets, it appears that the small Android market has shifted to behave more like the full market's R0 behavior than the full market has shifted to behave more like the small market. Meaning, in the case of exfiltration, the small Android market performs just as poorly as the large market, rather than the full market performs just as well as the smaller market.

The other exception is the interaction between the nonvendor market and the Apple and full Android market. In the case of malware spread, Apple-type behavior on the nonvendor market causes R0 to increase

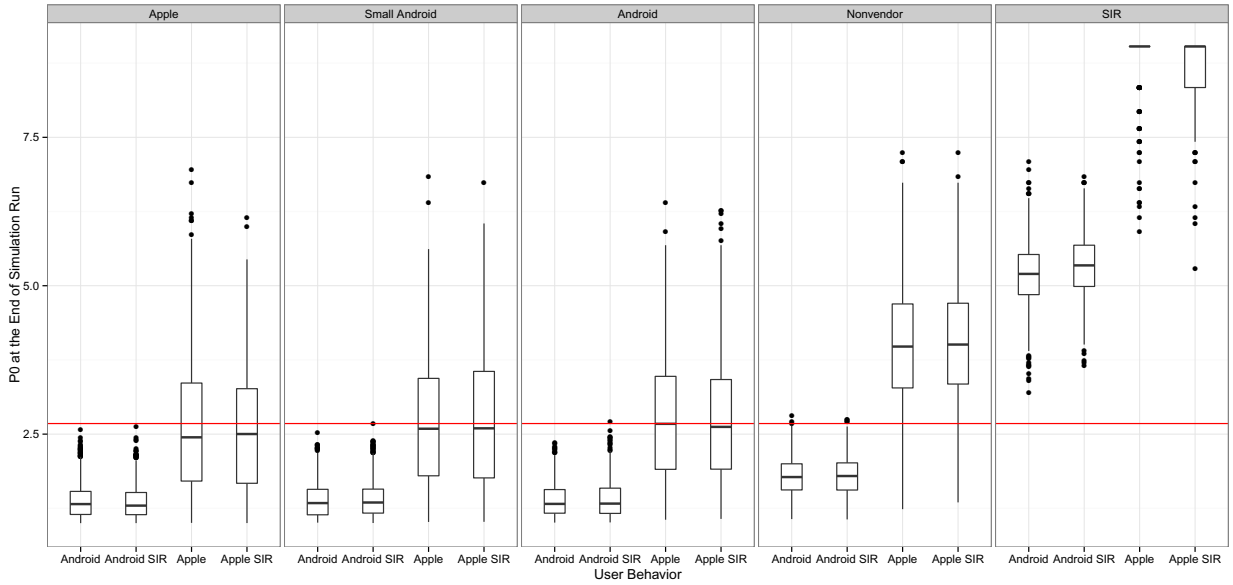
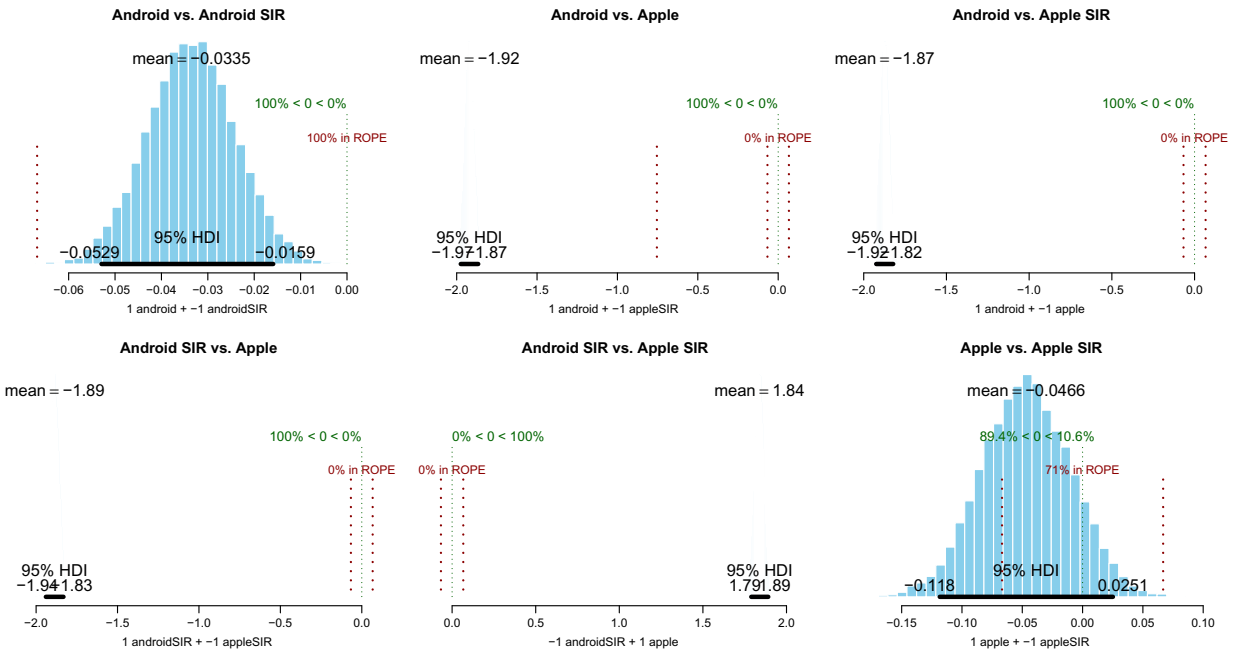


Figure 4.13: P0 at the end of Simulation Run by User Behavior and Market Structure. Markets are Ordered by Level of Centralized Control. Red Line Represents Mean of the Data.

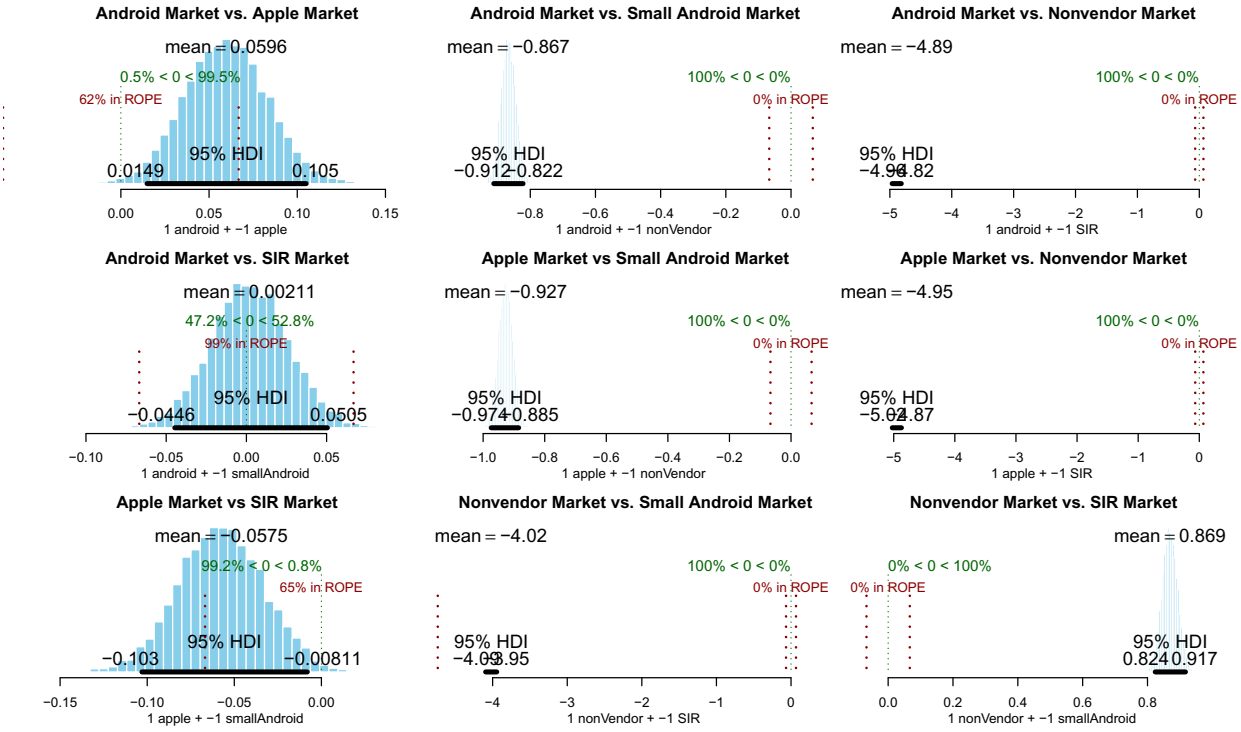
by 0.12 compared to Android-type behavior on the same market. This is a small, but credible increase in R0. When we look at information exfiltration, the same comparison leads to a 0.35 increase in P0. In all other interactions, however, the interaction effect is less intense.

This could be due to the fact that in the more controlled conditions, once the primary marketplace removes the malware, there is little additional spread and R0, leaving parts of the network revealed. However, due to the persistence of the malware in the nonvendor condition, more of the network is revealed by reaching more sources. The difference between the nonvendor and SIR conditions in regards to P0 is that, with the Apple user behavior, the entire network is infected in the SIR market condition, so there is no additional information revealed.

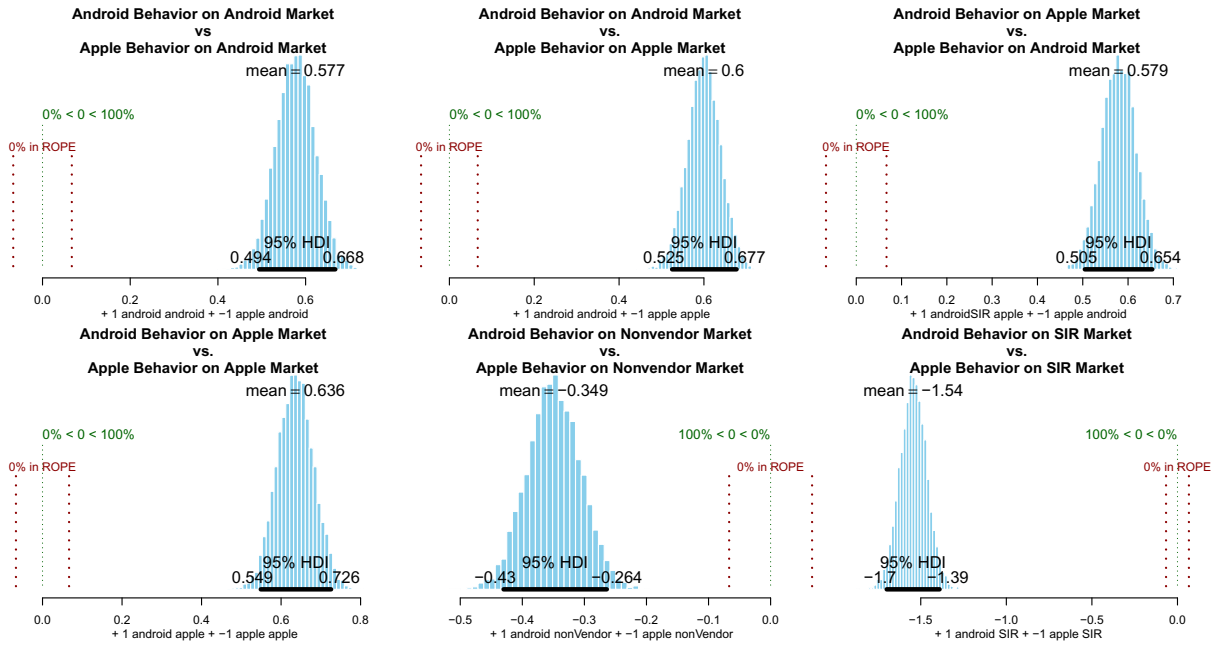
Our model shows that user behavior is a more important component of social malware spread than market structure in the case of number of devices infected and revealed. User behavior is just as important as market structure in determining infection duration, but it does not seem that infection duration is a useful metric for determining the severity of an outbreak. Our model also shows that information exfiltration as an attack vector is of limited use, if an attacker can fully compromise the network.



(a) Comparisons of the Effects of User Behavior on P0.



(b) Comparisons of the Effects of Market Structure on P0.



(c) Comparisons of Interaction Effects on P0.

Figure 4.14: Contrasts of Behavior (4.14a), Market (4.14b), and Interaction (4.14c) Effects on P0.

4.5 Discussion

Our model was designed to demonstrate the effectiveness of social pressure as a contagion vector, and some of the second-order effects of infection, such as information exfiltration. We manage to show that user behavior is an important component in social malware spread, and that market structures can limit potential dangers. We also show that, as the digital environment becomes less regulated, the more important individual behavior becomes.

However, our research fails to address certain aspects of user behavior. We only consider two types of users, those that follow a primarily local-based update scheme, Android users, and those that follow a well-balanced update scheme, Apple users. We do not know why those users have chosen those particular strategies, nor if those strategies hold for those users anymore. Furthermore, we also need to consider the type of user that would use a jail-broken iPhone. If users with jail-broken phones use a primarily local-based adoption strategy, we might be able to say that they and Android users are aware of potential risks and have adopted an effective strategy to mitigate those risks.

It might also be useful to identify differences in strategies based on sophistication. It may be that Apple users, in general, are less technically savvy, and thus use a strategy that places them at risk. However, it

may also be that they recognize that the Apple marketplace has made it difficult–though not impossible–for malware authors, so they need not be as concerned with the safety of the apps they choose to download. These are both outstanding questions that we could explore by further refining my models.

We also used several important assumptions in implementing our model. Primarily, we chose to use a weighting scheme based on reciprocal communication rather than using standard uniform selection and response. [149] This particular weighting scheme has important ramifications for information diffusion on networks. Our assumption essentially states that the network structure will change based on the cost of communication and the goal of communication. [140, 173] This assumption needs to be explored further, as we feel it can yield insight into differences in network structure, for network analysis itself, but also in terms of manufacturing of spam and phishing campaigns.

Our model has some important limitations in these regards. For example, every node in our network belongs to the same type of smart phone network. Additionally, we fold users' SMS and call networks into a single projection. This simplifies the model, but may not represent the effects of communication investment accurately. For example, Twitter-type networks may have a certain structure because they use broadcast communication, and so can draw attention to, but are unlikely to modify opinions. [85, 121, 150, 151] On the other hand, use of direct communication, such as telephones or face-to-face communication, while requiring more resources, is more likely to sway opinion. [3]

Our model also has some use in terms of studying risks of friendship, structure of the network, and contagion. [4, 223, 224] We collect a full state space matrix, not only of individual infections and data loss, but also which neighbors and markets lead to infection. We have not the time to fully analyze the full state space and the interactions between network structure and the spread of malware with the weighting scheme we used, but we are interested in examining how this weighting affects optimal structures for diffusion. [155, 223]

4.6 Conclusions

As mobile devices have become more ubiquitous, they have attracted the attention of malware developers. The interface between technical security measures and human behavior, in the form of app permission granting, does not work. Leaders in mobile operating systems have implemented other ways to limit users' exposure to malware, but they are not perfect. Users must make decisions about which applications they

install and use on their mobile devices with very limited knowledge about those applications' true functionality.

Users make these decisions based on a number of factors, primarily relying on social cues, such as their social network or the number of download counts at an app store. Malware authors have already taken advantage of these social cues to entice users to download and install mobile malware. We examined users' app search and adoption strategies interact with different market structures to see how the affect spread of malware and information exfiltration.

We had two key results from our simulations. First, user behavior was the key factor in limiting the spread of malicious apps on markets. In particular, when our agents rely more heavily on close contacts, rather than download counts, it slows the propagation of malware. This social behavior becomes more important as regulation over the market place decreases. The strategy of relying on social networks to filter applications relates back to our vigilance parameter η, from Chapter 3.

Secondly, and unsurprisingly, centralized markets are the best at reducing the spread of malware. More importantly, however, is that centralized markets with access to third-party market places, are not significantly more effective at reducing malware in comparison to a single, monolithic market place. So, while a single, centralized market can limit risky user behavior, it is not significantly better at limiting malware spread from a more open market place. The fact that open markets, with a primary central market, offer just as much control on the propagation of mobile apps has important design ramifications.

In particular, this suggests that having a well-informed, or if we consider Chapter 3, a risk-averse user base can be effective at limiting the effects of a given outbreak. It also demonstrates the usefulness in having a primary, centralized market for less informed users. The model also shows that, outside of the centralized market, third party market places can also thrive without significantly affecting the likelihood of malware propagation. This ties into the results from Chapter 2, in that these results require some sort of differentiation based on ability and behavior. The model in this chapter demonstrates, using observed preferences in mobile device users, that this is the case. The full extent of the effects of user behaviors on propagation of malware remains to be seen, but, by tying observed behaviors with large-scale models, we have a better picture of the potential.

5 Conclusions

My work has focused on trying to identify patterns in digital security behaviors and reason about the large scale effects of those behaviors. In the first study I examined attention to security cues in web browsers using eye tracking. Next I studied a model focused the effects of interactions between risk-averse and risk-taking populations for large-scale botnets. Finally, I evaluated a model of mobile malware spread based on dominate modes of app adoption strategies and marketplace implementations. All of these studies attempt to examine how users participate in socio-technical systems and how individual behaviors in the aggregate create or exacerbate systemic vulnerabilities.

In our eye tracking study we found that expertise was only a small factor in attention to security cues. Task-type was a much bigger indicator of attention to cues, with tasks requiring the use of personal accounts driving attention to cues. In fact, in most tasks experts performed the same as novices. Additionally, novices did pay attention to security cues, which suggests a more complicated behavior than previous studies in user attention to security. Our results give us a glimpse into how the security socio-technical system interacts with user behavior and sophistication.

In the second study I examined how differences in security awareness might affect global malware prevalence. I introduced and evaluated a coupled two-population epidemiological model that incorporated user vigilance, cost for risk-averse behavior, and social response. We found that cost is the largest factor for affecting malware prevalence, outside of infection rates of malware. While we found that individual recovery was an important aspect in malware mitigation, the likelihood that individuals can discover and recover from malware at the same rate malware spreads, limits its real-world importance. When that is taken into account, social response becomes very important for mitigating malware spread.

Finally, in our third study, I looked at how different behaviors of mobile app adoption and market place implementations affect the spread of malware via social channels. We found that reliance on local connections was an effective mitigator for social malware spread particularly in unregulated markets. Local

app discovery also limited the rate of information exfiltration in a network of mobile phone users. Well-regulated markets are able to limit the effects of effects of more risky behaviors. Without the regulation of official marketplaces, reliance on unknown, global influence leads to large increases in the rate of spread of malware.

By focusing on task context and sophistication, I present a more nuanced view of users' participation in available security technologies. Then, by examining two populations (i.e., risk-takers and risk-averse) I look at how costs and social interactions can affect malware spread in aggregate. Finally, I focus on how users of different mobile devices utilize different strategies for app adoption and show that those strategies have important ramifications for the spread of mobile malware.

Each study examines ways in which users utilize their technology, the aggregate effects of those behaviors, or both. We can use empirical behavioral studies to develop a better understanding of digital behavior and present a more holistic approach to information security. We then use models to extrapolate potential large-scale implications of those digital behaviors. The combination of user behavior studies and large scale models give us a more accurate picture of potential risks and, when coupled with economic data, a better idea of true loses.

5.1 Future Work

My work is only a limited part of a broader attempt to address the complicated issues of information security with multidisciplinary tools and techniques. My own work suggests important future study into the nature of expertise, task-context, and environmental factors on users' behavior. In particular, initial pilot studies using our eye-tracking methodology suggest significant cultural factors in the way users approach privacy-preserving behavior. In regards to my modeling work, we are seeking additional data so we can validate our models against full data sets. My work in mobile malware suggests work on the effects of multi-networks on malware spread. There are also issues regarding proper weighting of our networks in the simulations, as well as what the different strategies users have employed represent.

In the following section I will address possible future work based on my previous studies. I will first focus on results of two pilot studies that suggest, perhaps obviously, the importance of culture and technological familiarity in assessing security behavior. I will follow that by looking at the difficulty of validating our models, but point out several pieces of data that seem to suggest that we are moving in the right di-

rection. Finally, I will look at ways to refine my models, such as looking at the difference between Apple device users that have not jail-broken their devices and those that have and integrating risk-aversion into my agent-based model.

Eye Tracking

Our eye tracking study demonstrated that task-context was a bigger indicator of visual attention to security cues than expertise. However, we left a lot of data left unanalyzed. We only looked at the number of fixations and the dwell time/fixation. There are several other aspects that we need to examine, such as likelihood of arriving at the area of interest randomly, compared with number of times users actually fixated within the time frame. There may be other areas of the screen that users are utilizing to make security decisions.

We also need to analyze the importance of potential loss. In our experiment, we manipulated whether or not participants logins were sent encrypted or in cleartext. However, during the cleartext manipulation, none of the participants used the security cues. None of the participants were using their own OpenID identity providers either. We also know that two of the participants did not use their own accounts during the encrypted manipulation. The behavior during the OpenID portion suggests that the potential for personal loss affects attention to security cues. However, during the experiment, we did not record which subjects used their own accounts. This becomes more important when we look at a pilot study we ran with Japanese participants.

When we ran our study with Australian participants, all but two participants used their own Facebook accounts to participate in the study. However, when we ran a pilot study with Japanese participants, all but three of the Japanese participants used our lab-generated credentials. There is research that suggests that Japanese participants perceive risk differently than Western participants. [125] Understanding why this is the case requires a cognitive-experimental model of digital behavior. [212]

Model Validation

In my modeling work I proposed two different models that we used to reason about aspects affecting malware spread. In both of the models we included aspects of social mitigation of malware spread. In the first model, described in Chapter 3, I looked at effects of risk-aversion on malware spread. While we were able to fit the the model to phishing data, we lack data on users' risk-aversion. We also made several assumptions about the the connection between phishing sites and malware sites, and the long-term prevalence of

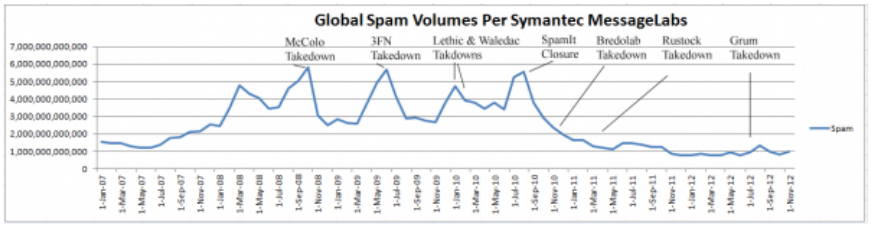


Figure 5.1: Stable Levels of Spam from January 1, 2007 through November 1, 2012 [124]

malware. We were unable to find data on risk-aversion, but several reports by Google and Microsoft seem to support our other assumptions.

Google maintains an up-to-date tally of compromised machines they monitor. As part of their report they publish the number of phishing and malware sites they detect. In this report they demonstrate a fair correlation between phishing and malware sites. For our model, we had data on phishing websites and made the assumption that phishing websites were proportional to malware sites, and Google's transparency report seems to support this assumption. [89] The Google report, which was published one year after our publication, along with data about phishing messages (Figure 5.1) , also confirmed our assumption that we could model as a SIS process, due to the relatively stable level of malware across time. [89, 124]

Similarly, we found a Microsoft report on reputation management in limiting the spread of malware. The Microsoft report does not give specific numbers for malware downloads, they do provide a download traffic/hour curve (Figure 5.2). They show that the reputation-based alert system is able to respond more rapidly than traditional anti-virus signatures, confirming our hypothesis about the effectiveness of social-response and risk-communication. [96]

A secondary Microsoft report, published in March of this year, also confirmed our assumption about the nature of risk-averse entities. We assumed that risk-averse entities were less likely to become infected, but were not perfectly protected. Microsoft reported that the software that they use to find the right program to run an unknown file type has been modified, via update servers, to install malware. [136] It took Microsoft approximately a week to discover the connection and remove the compromised version of the software, but new devices were being compromised for roughly a month after that point (Figure 5.3).

In my second model, described in Chapter 4, we looked at user behavior and market implementation in the spread of malware on mobile devices. While we had data on user behavior to parameterize our model, we lacked information on specific malware spread and amount of malware on 3rd party marketplaces. While we have been unable to find dynamics data for mobile malware, F-Secure published a report in the summer

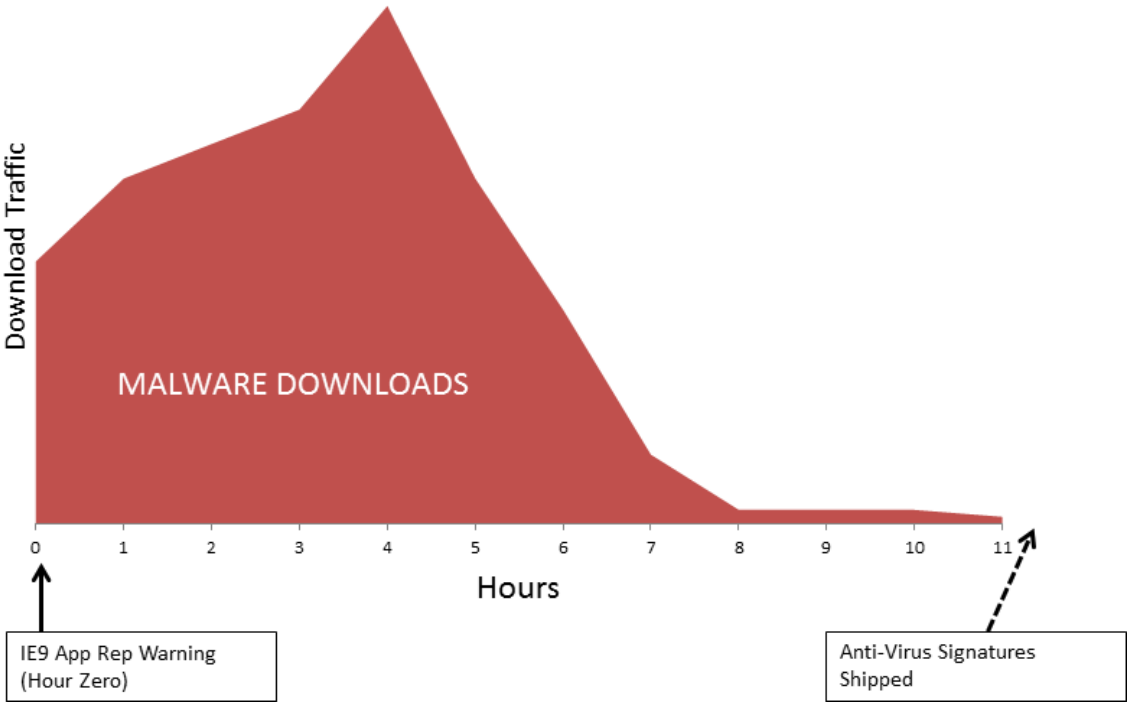


Figure 5.2: Effects of Reputation-Based Warning on Malware Spread [96]

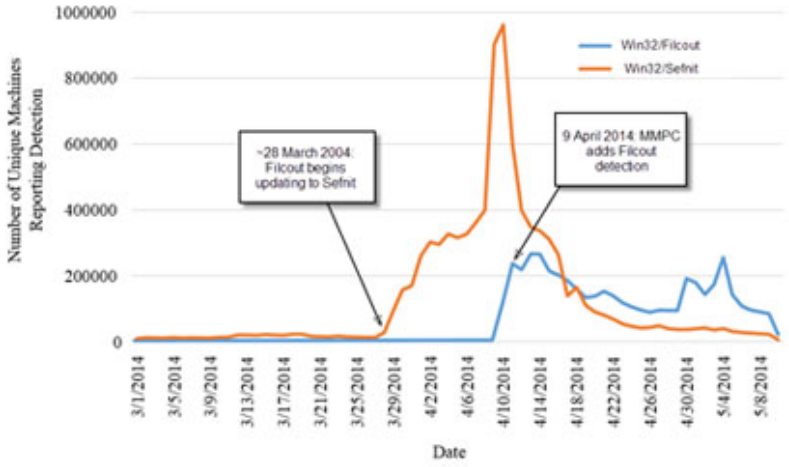


Figure 5.3: Downloads of Microsoft's Compromised Software Detection Software [136]

of 2014–between writing and publication of this work–that supports our assumptions about the structure and amount of malware on 3rd-party Android app markets.

More recently, as I was finalizing my modeling work, significant protests broke out in Hong Kong. The protesters were using their mobile devices for communication. Multiple versions of the apps being used to communicate were infected with malware that recorded and offloaded contact and communication information, exactly like our proposed malware. This particular situation matches our non-Vendor market place condition, as the mobile market place is comprised of only third-party vendors. In this case, due to social connections and need for shared applications, the mobile malware spread rapidly through the protesters, matching the behavior predicted by my model. [152]

The discovery of the Microsoft report on reputation-based detection and removal helps justify our inclusion of social response in our mobile malware model as well. [96] While we were unable to find specific malicious app adoption curves, we were able to find some graphical illustrations of popular apps, such as Angry Birds. [131] We know our model can capture the dynamics found in Figure 5.4. We expect that mobile malware, particularly malware that is good at hiding, will have similar adoption curves, but, as the malicious nature is discovered, users will rapidly uninstall the app.

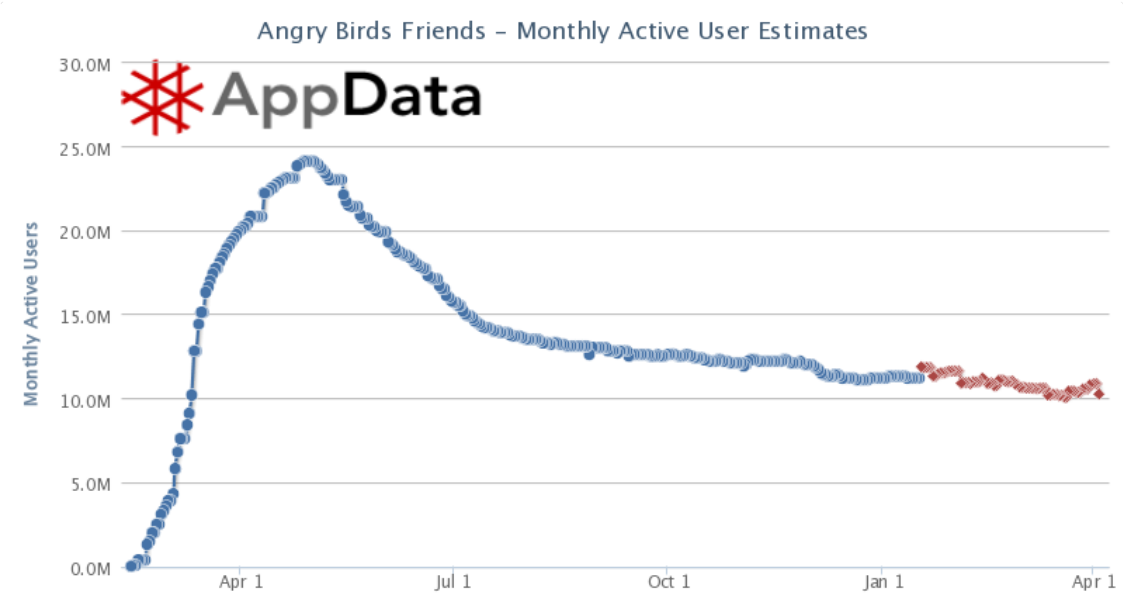


Figure 5.4: Downloads of Rovio's Angry Birds Friends [131]

Finally, a report by F-Secure reveals the structure of the Android marketplace and an approximate level of malware prevalence on the marketplace. In our model we used number of apps as a proxy for the size of

a market and assumed a long-tailed distribution of 3rd-party marketplaces. F-Secure's report confirms this, showing the Google Play store dominating the markets, and the 3rd-party marketplaces following a long-tailed distribution of apps hosted. We also assumed that the smaller marketplaces would be less responsive to malware on the market. This does not seem to be the case, though 3rd-party marketplaces are less responsive than the official store. [192]

While some aspects of our models seem to have been validated, there are still aspects that require additional data. While we can fit our model of risk-aversion to many datasets, this is largely meaningless until we actually have information about participants risk-aversion. Similarly, we have reasonable parameters for user behavior and marketplace structure and behavior, but we have no good data about the time-scales of mobile malware infections. The most invasive malware (i.e., malware that makes itself known in obvious ways) is caught quickly. In the case of Bill Shocker, this was at the end of the month, when cell phone bills came due, but we have no idea about more subtle malware. [179]

Model Refinement

Both of the models presented can use additional refinements, or possibly a way to combine the two. The first model is a mathematical model, which allows for higher level reasoning, but lacks the fine-grained, processed-based approach of the second model. Yet, the parameters can be translated fairly easily. For example, rather than being a two-population model, each agent gets a level of risk-aversion that can be updated by the behavior and infection status of it neighbors. Similarly, the cost of risk-aversion can be made to control the dynamics of risk-aversion updates.

There are also aspects of the second model that need to be considered, particularly in light of work in multi-networks. In the second model, I collapsed two networks (a call network and a SMS network) into a single network. However, recent work regarding communication on network suggests that the type of network affects the weight given to the communication. [140] A further refinement of the second model would be to separate the networks and investigate the effects that mode of communication has on the malware spread.

Additionally, the second model only considers two types of users, those that follow a primarily local-based update scheme, Android users, and those that follow a well-balanced update scheme, Apple users. We do not know why those users have chosen those particular strategies, nor if those strategies hold for those users anymore. It would be useful to investigate what effects using a jail-broken iPhone might have on app

adoption strategy. If users with jail-broken phones use a primarily local-based adoption strategy, we might be able to say that they, and Android users are aware of potential risks and have adopted an effect strategy of mitigating those risks.

It might also be useful to identify differences in strategies based on sophistication. It may be that Apple users, in general, are less technically savvy, and thus use a strategy that places them at risk. However, it may also be that they recognize that the Apple marketplace has made it difficult–though not impossible–for malware authors, so they need not be as concerned with the safety of the apps they choose to download. These are both outstanding questions that we could explore by further refining my models.

5.2 Concluding Remarks

I have presented a few studies that attempt to answer some of the myriad questions we have about cognition and information security. However, these are not the only methods. I have demonstrated that information security can benefit from drawing on techniques in cognitive science and complex systems to approach the questions we have from a bottom-up perspective. The bottom-up perspective can be compared with a top-down approach to find where there is agreement, and where there is disagreement, we can direct more intensive research.

While quantitative methods and behavior research are useful for clearly defining and exploring a problem space, information security has a lot to gain from less quantitative, but no less rigorous, methods found in the social sciences and philosophy. More qualitative research can be used to better map the problem space before and to help develop more rigorous quantitative methodologies. Furthermore, information security has important contributions to add to other fields. For example, continuing the tradition of early studies in network effects and epidemiology in information security, my second model explicitly changed the uniform selection of neighbors in network spread models in an attempt to capture the nature of trust relationships inherent in a communication network. [113, 114, 157]

Information security also has important insights on how users perceive and make decisions in uncertain digital environments. The psychology of risk and decision-making in digital environments is largely unexplored. However, exploring the factors that impact risk perception and decision-making requires both an in-depth knowledge of psychological and cognitive science research methodologies and an understanding of the relevant aspects of information security. In particular, the cultural effects on forming risk perception

and responding to potential risk presents an interesting problem for developing cross-cultural digital risk communication.

My work has demonstrated the effectiveness of a robust methodology in predicting widespread malware propagation. While, it is only a small part of the information security picture, it suggests particularly useful extensions. For example, with applicable economic data, we can use my models to extrapolate realistic global economic loss. By inferring certain network behaviors, we can look at the extent of privacy loss due to malware infection. This particular inference can be helpful in gauging the extent of repressive government surveillance in given communities, such as in the Hong Kong protests.

My work develops a methodology that fully explores the ecology of security. This ecology of security must include studying how users manipulate their devices and digital environments. It must also examine how those environments, in turn, affect how users perceive and interact in those same complex and uncertain habitats. But, the better we understand the various interactions between people, their environments (both physical and digitally contructed), and their behavior the better equipped we will be to respond to the ramifications of those interactions.

Bibliography

[1] The OAuth 1.0 Protocol, Apr. 2010.

[2] A. Acquisti and R. Gross. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *Priv. Enhancing Technol.*, pages 36–58. Springer, 2006.

[3] N. Aharony, W. Pan, C. Ip, I. Khayal, and A. Pentland. Social fMRI: Investigating and shaping social mechanisms in the real world. *Pervasive Mob. Comput.*, 7(6):643–659, Dec. 2011.

[4] C. G. Akcora, B. Carminati, and E. Ferrari. Risks of Friendships on Social Networks. *arXiv Prepr. arXiv1210.3234*, (Icdm):10, Oct. 2012.

[5] M. E. Alexander and R. Kobes. Effects of vaccination and population structure on influenza epidemic spread in the presence of two circulating strains. *BMC Public Health*, 11 Suppl 1(Suppl 1):S8, Jan. 2011.

[6] R. Anderson. Why information security is hard-an economic perspective. In *Comput. Secur. Appl. Conf. 2001. ACSAC 2001. Proc. 17th Annu.*, pages 358–365. IEEE, 2001.

[7] R. Anderson and T. Moore. The economics of information security. *Science*, 314(5799):610–3, Oct. 2006.

[8] R. Anderson and T. Moore. Information security: where computer science, economics and psychology meet. *Philos. Trans. A. Math. Phys. Eng. Sci.*, 367(1898):2717–27, July 2009.

[9] J.-B. André, J.-B. Ferdy, and B. Godelle. Within-host parasite dynamics, emerging trade-off, and evolution of virulence with immune system. *Evolution*, 57(7):1489–97, July 2003.

[10] Anti-Phishing Working Group. Phishing Activity Trends Report 2nd Half 2010. Technical Report December, Anti-Phishing Working Group, 2010.

[11] Anti-Phishing Working Group. Phishing Activity Trends Report 2nd Quarter 2010. Technical Report June, 2010.

[12] Anti-Phishing Working Group. Phishing Activity Trends Report 1st Half 2011. Technical Report June, Anti-Phishing Working Group, 2011.

[13] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster. Building a dynamic reputation system for DNS. In *19th Usenix Secur. Symp.*, 2010.

[14] Apple Inc. App Review, 2014.

[15] Apple Inc. Common App Rejections, 2014.

[16] M. Arianezhad, D. Stebila, and B. Mozaffari. Usability and Security of Gaze-Based Graphical Grid Passwords. In A. Adams, M. Brenner, and M. Smith, editors, *Financ. Cryptogr. Data Secur. SE - 2*, volume 7862 of *Lecture Notes in Computer Science*, pages 17–33. Springer Berlin Heidelberg, 2013.

[17] F. Asgharpour, D. Liu, and L. J. Camp. Mental models of computer security risks. *Work. Econ. Inf. Secur.*, 2007.

[18] T. August and T. I. Tunca. Let the Pirates Patch? An Economic Analysis of Software Security Patch Restrictions. *Inf. Syst. Res.*, 19(1):48–70, Mar. 2008.

[19] R. F. Baggaley, G. P. Garnett, and N. M. Ferguson. Modelling the Impact of Antiretroviral Use in Resource-Poor Settings. *PLoS Med.*, 3(4):e124, 2006.

[20] M. Bailey, E. Cooke, F. Jahanian, D. Watson, and J. Nazario. The Blaster Worm: Then and Now. *IEEE Secur. Priv. Mag.*, 3(4):26–31, July 2005.

[21] D. Baker. The Evolved Threat Paradigm: Look Who's Wearing the Black Hats! In *Proc. 1992-1993 Work. New Secur. Paradig.*, pages 126–130. ACM, 1993.

[22] A.-L. Barabási, R. Albert, and H. Jeong. Scale-free characteristics of random networks: the topology of the world-wide web. *Phys. A Stat. Mech. its Appl.*, 281(1-4):69–77, June 2000.

[23] L. Bauer, C. Bravo-Lillo, E. Fragkaki, and W. Melicher. A comparison of users' perceptions of and willingness to use Google, Facebook, and Google+ single-sign-on functionality. *Proc. 2013 ACM Work. Digit. identity Manag. - DIM '13*, pages 25–36, 2013.

[24] D. Bell. Secure computer systems: Mathematical foundations. Technical report, DTIC Document, 1973.

[25] N. M. Bellows, B. W. Bellows, and C. Warren. Systematic Review: the use of vouchers for reproductive health services in developing countries: systematic review. *Trop. Med. Int. Heal.*, 16(1):84–96, Jan. 2011.

[26] K. Benton, L. J. Camp, and V. Garg. Studying the effectiveness of android application permissions requests. In *2013 IEEE Int. Conf. Pervasive Comput. Commun. Work. (PERCOM Work.*, number March, pages 291–296. IEEE, Mar. 2013.

[27] D. Besnard and B. Arief. Computer security impaired by legitimate users. *Comput. Secur.*, 23(3):253–264, May 2004.

[28] L. Bilge and T. Dumitras. Before we knew it: an empirical study of zero-day attacks in the real world. In *Proc. 2012 ACM Conf. Comput. Commun. Secur. - CCS '12*, page 833, New York, New York, USA, 2012. ACM Press.

[29] S. Blower and H. Dowlatabadi. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *Int. Stat. Rev.*, 62(2), 1994.

[30] R. Böhme and S. Köpsell. Trained to accept? In *Proc. 28th Int. Conf. Hum. factors Comput. Syst. - CHI '10*, page 2403, New York, New York, USA, 2010. ACM Press.

[31] M.-C. Boily, F. I. Bastos, K. Desai, and B. Mâsse. Changes in the transmission dynamics of the HIV epidemic after the wide-scale use of antiretroviral therapy could explain increases in sexually transmitted infections: results from mathematical models. *Sex. Transm. Dis.*, 31(2):100–13, Feb. 2004.

[32] J. Bonneau and S. Preibusch. The privacy jungle: On the market for data protection in social networks. In *Econ. Inf. Secur. Priv.*, pages 121–167. Springer, 2010.

[33] R. Brady, R. Anderson, and R. C. Ball. Murphy's law, the fitness of evolving species, and the limits of software reliability. Aug. 1999.

[34] C. Bravo-Lillo and L. Cranor. Improving computer security dialogs. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, editors, *Human-Computer Interact. – INTERACT 2011*, pages 18–35. Springer Berlin Heidelberg, 2011.

[35] C. Bravo-Lillo, L. F. Cranor, J. Downs, and S. Komanduri. Bridging the Gap in Computer Security Warnings: A Mental Model Approach. *IEEE Secur. Priv. Mag.*, 9(2):18–26, Mar. 2011.

[36] C. Bravo-Lillo, S. Komanduri, L. F. Cranor, R. W. Reeder, M. Sleeper, J. Downs, and S. Schechter. Your attention please. In *Proc. Ninth Symp. Usable Priv. Secur. - SOUPS '13*, page 1, New York, New York, USA, 2013. ACM Press.

[37] H. M. Burke, K. F. Pedersen, and N. E. Williamson. An assessment of cost, quality and outcomes for five HIV prevention youth peer education programs in Zambia. *Health Educ. Res.*, 27(2):359–69, Apr. 2012.

[38] M. Burns. Eric Schmidt: "There Are Now 1.3 Million Android Device Activations Per Day", 2012.

[39] L. Camp. Mental models of privacy and security. *IEEE Technol. Soc. Mag.*, 28(3):37–46, Jan. 2009.

[40] L. Camp and C. Wolfram. Pricing security. In *Proc. CERT Inf. Surviv. Work.*, pages 31–39. Citeseer, 2000.

[41] D. Chakrabarti, C. Faloutsos, Y. Wang, and C. Wang. Epidemic spreading in real networks: an eigenvalue viewpoint. In *22nd Int. Symp. Reliab. Distrib. Syst. Proceedings.*, number 22, pages 25–34. IEEE Comput. Soc, 2003.

[42] S. Chhabra, A. Aggarwal, F. Benevenuto, and P. Kumaraguru. Phi.sh/$oCiaL. In *Proc. 8th Annu. Collab. Electron. Messag. Anti-Abuse Spam Conf. - CEAS '11*, pages 92–101, New York, New York, USA, 2011. ACM Press.

[43] S. Chiasson, A. Forget, R. Biddle, and P. C. Oorschot. User interface design affects security: patterns in click-based graphical passwords. *Int. J. Inf. Secur.*, 8(6):387–398, May 2009.

[44] E. Chin, A. P. Felt, V. Sekar, and D. Wagner. Measuring user confidence in smartphone security and privacy. *Proc. Eighth Symp. Usable Priv. Secur. - SOUPS '12*, (1):1, 2012.

[45] J. Choi, C. Fershtman, and N. Gandal. Network Security: Vulnerabilities and Disclosure Policy. 2007.

[46] S. Choney. Apple App Store infiltrated by researchers' 'Jekyll' malware, 2014.

[47] N. a. Christakis and J. H. Fowler. The collective dynamics of smoking in a large social network. *N. Engl. J. Med.*, 358(21):2249–58, May 2008.

[48] N. Christin, S. Egelman, T. Vidas, and J. Grossklags. It's All about the Benjamins: An Empirical Study on Incentivizing Users to Ignore Security Advice. In G. Danezis, editor, *Financ. Cryptogr. Data Secur.*, volume 7035 of *Lecture Notes in Computer Science*, pages 16–30. Springer Berlin Heidelberg, 2012.

[49] D. Clark and D. Wilson. A Comparison of Commercial and Military Computer Security Policies. In *Proc. IEEE Symp. Secur. Priv.*, pages 184–194. Published by the IEEE Computer Society, 1987.

[50] M. D. Conover, C. Davis, E. Ferrara, K. McKelvey, F. Menczer, and A. Flammini. The geospatial characteristics of a social movement communication network. *PLoS One*, 8(3):e55957, Jan. 2013.

[51] M. D. Conover, B. Gonçalves, A. Flammini, and F. Menczer. Partisan asymmetries in online political activity. *EPJ Data Sci.*, 1(1):6, 2012.

[52] L. Constantin. Security Experts Concerned About Google's Attitude Toward Android Malware, 2011.

[53] D. Dagon, G. Gu, C. P. Lee, and W. Lee. A Taxonomy of Botnet Structures. In *Twenty-Third Annu. Comput. Secur. Appl. Conf. (ACSAC 2007)*, pages 325–339. IEEE, Dec. 2007.

[54] D. Dagon, C. Zou, and W. Lee. Modeling botnet propagation using time zones. In *Proc. 13th Annu. Netw. Distrib. Syst. Secur. Symp.* Citeseer, 2006.

[55] F. Darabi Sahneh and C. Scoglio. Epidemic spread in human networks. In *IEEE Conf. Decis. Control Eur. Control Conf.*, pages 3008–3013. IEEE, Dec. 2011.

[56] S. B. Day and R. L. Goldstone. Analogical transfer from a simulated physical system. *J. Exp. Psychol. Learn. Mem. Cogn.*, 37(3):551–67, May 2011.

[57] D. E. Denning. Secure personal computing in an insecure network. *Commun. ACM*, 22(8):476–482, Aug. 1979.

[58] D. E. Denning. A new paradigm for trusted systems. In *Proc. 1992-1993 Work. New Secur. Paradig. - NSPW '92-93*, pages 36–41, New York, New York, USA, 1993. ACM Press.

[59] D. E. Denning and P. J. Denning. Data Security. *ACM Comput. Surv.*, 11(3):227–249, Sept. 1979.

[60] P. J. Denning. Editorial: what is software quality? *Commun. ACM*, 35(1):13–15, Jan. 1992.

[61] R. Dhamija, J. D. Tygar, and M. Hearst. Why phishing works. In *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI) 2006*, pages 581–590. ACM, 2006.

[62] J. S. Downs, M. Holbrook, and L. F. Cranor. Behavioral response to phishing risk. In *Proc. anti-phishing Work. groups 2nd Annu. eCrime Res. summit - eCrime '07*, pages 37–44, New York, New York, USA, 2007. ACM Press.

[63] M. Dürmuth, T. Güneysu, and M. Kasper. Evaluation of standardized password-based key derivation against parallel processing platforms. In S. Foresti, M. Yung, and F. Martinelli, editors, *Comput. Secur. – ESORICS 2012*, pages 716–733. Springer Berlin Heidelberg, 2012.

[64] B. Edelman. Adverse selection in online "trust" certifications. In *Proc. 11th Int. Conf. Electron. Commer. - ICEC '09*, page 205, New York, New York, USA, 2009. ACM Press.

[65] B. Edwards, T. Moore, G. Stelle, S. Hofmeyr, and S. Forrest. Beyond the blacklist. In *Proc. 2012 Work. New Secur. Paradig. - NSPW '12*, page 53, New York, New York, USA, 2012. ACM Press.

[66] S. Egelman. *Trust Me: Design Patterns for Constructing Trustworthy Trust Indicators*. PhD thesis, Carnegie Mellon University, April 2009.

[67] S. Egelman, A. Felt, and D. Wagner. Choice Architecture and Smartphone Privacy: There's A Price For That. In *Work. Econ. Inf. Secur.*, Berlin, Germany, 2012.

[68] Electronic Frontier Foundation. The EFF SSL Observatory, 2010.

[69] W. Enck, D. Octeau, P. McDaniel, and S. Chaudhuri. A study of android application security. In *Proceeding 20th USENIX Secur. Symp.*, number August, 2011.

[70] W. Enck, M. Ongtang, and P. McDaniel. On lightweight mobile phone application certification. In *Proc. 16th ACM Conf. Comput. Commun. Secur.* ACM, 2009.

[71] A. Felt, S. Egelman, M. Finifter, D. Akhawe, and D. Wagner. How to ask for permission. In *Proc. 7th USENIX Conf. Hot Top. Secur.* USENIX Association, 2012.

[72] A. P. Felt, M. Finifter, E. Chin, S. Hanna, and D. Wagner. A survey of mobile malware in the wild. In *Proc. 1st ACM Work. Secur. Priv. smartphones Mob. devices - SPSM '11*, page 3, New York, New York, USA, 2011. ACM Press.

[73] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner. Android Permissions: User Attention, Comprehension, and Behavior. In *Proc. Eighth Symp. Usable Priv. Secur. - SOUPS '12*, page 1, New York, New York, USA, 2012. ACM Press.

[74] N. Ferguson and G. Garnett. More realistic models of sexually transmitted disease transmission dynamics: Sexual partnership networks, pair models, and moment closure. *Sex. Transm. Dis.*, 27(10):1–10, 2000.

[75] A. Forget, S. Chiasson, and R. Biddle. Shoulder-surfing resistance with eye-gaze entry in cued-recall graphical passwords. In *Proc. 28th Int. Conf. Hum. factors Comput. Syst. - CHI '10*, page 1107, New York, New York, USA, 2010. ACM Press.

[76] S. Fortunato, A. Flammini, and F. Menczer. Scale-Free Network Growth by Ranking. *Phys. Rev. Lett.*, 96(21):218701, May 2006.

[77] L. Franzini, E. Marks, P. F. Cromwell, J. Risser, L. McGill, C. Markham, B. Selwyn, and C. Shapiro. Projected economic costs due to health consequences of teenagers' loss of confidentiality in obtaining reproductive health care services in Texas. *Arch. Pediatr. Adolesc. Med.*, 158(12):1140–6, Dec. 2004.

[78] B. Friedman, D. Hurley, D. C. Howe, E. Felten, and H. Nissenbaum. Users' conceptions of web security. In *CHI '02 Ext. Abstr. Hum. factors Comput. Syst. - CHI '02*, page 746, New York, New York, USA, 2002. ACM Press.

[79] S. Furnell. Why users cannot use security. *Comput. Secur.*, 24(4):274–279, June 2005.

[80] a. Ganesh, L. Massoulie, and D. Towsley. The effect of network topology on the spread of epidemics. In *Proc. IEEE 24th Annu. Jt. Conf. IEEE Comput. Commun. Soc.*, pages 1455–1466. IEEE, 2005.

[81] V. Garg and J. Camp. Heuristics and Biases: Implications for Security Design. *IEEE Technol. Soc. Mag.*, 32(1):73–79, Jan. 2013.

[82] V. Garg and L. Camp. Risk communication design: video vs. text. *Priv. Enhancing ...*, 2012.

[83] V. Garg, T. Koster, and L. Camp. Cross-country analysis of spambots. *EURASIP J. Inf. Secur.*, 2013(1):3, 2013.

[84] S. Gaw and E. W. Felten. Password management strategies for online accounts. *Proc. Second Symp. Usable Priv. Secur. - SOUPS '06*, page 44, 2006.

[85] J. Ghaderi and R. Srikant. Opinion Dynamics in Social Networks: A Local Interaction Game with Stubborn Agents. pages 1–35, Aug. 2012.

[86] R. L. Goldstone and S. B. Day. Introduction to "New Conceptualizations of Transfer of Learning". *Educ. Psychol.*, 47(3):149–152, July 2012.

[87] Google Corporation. Google Play Developer Policy.

[88] Google Corporation. Google Play Terms of Service.

[89] Google Corporation. Safe Browsing, 2014.

[90] S. Grabner-Kräuter. The Role of Consumers' Trust in Online-Shopping. *J. Bus. Ethics*, 39(1 - 2):43–50, 2002.

[91] S. Grabner-Kräuter and E. a. Kaluscha. Empirical research in on-line trust: a review and critical assessment. *Int. J. Hum. Comput. Stud.*, 58(6):783–812, June 2003.

[92] J. Gray, R. Klefstad, and M. Mernik. Adaptive and evolvable software systems: techniques, tools, and applications. In *37th Annu. Hawaii Int. Conf. Syst. Sci. 2004. Proc.*, page 1 pp. IEEE, 2004.

[93] R. T. Gray, K. W. Beagley, P. Timms, and D. P. Wilson. Modeling the Impact of Potential Vaccines on Epidemics of Sexually Transmitted Chlamydia trachomatis Infection. *J. Infect. Dis.*, 199(11):1680–1688, June 2009.

[94] C. Grier, K. Thomas, V. Paxson, and M. Zhang. @spam. In *Proc. 17th ACM Conf. Comput. Commun. Secur. - CCS '10*, page 27, New York, New York, USA, 2010. ACM Press.

[95] S. Gujrati and E. Y. Vasserman. The usability of truecrypt, or how i learned to stop whining and fix an interface. In *Proc. third ACM Conf. Data Appl. Secur. Priv. - CODASPY '13*, page 83, New York, New York, USA, 2013. ACM Press.

[96] J. Haber. SmartScreen Application Reputation in IE9, 2011.

[97] K. Hamacher and S. Katzenbeisser. Public security: Simulations Need to Replace Conventional Wisdom. In *Proc. 2011 Work. New Secur. Paradig. Work. - NSPW '11*, page 115, New York, New York, USA, 2011. ACM Press.

[98] E. Hargittai, L. Fullerton, E. Menchen-Trevino, and K. Y. Thomas. Trust online: Young adults' evaluation of web content. *Int. J. . . .*, 4:468–494, 2010.

[99] S. Helleringer and H. P. Kohler. Social networks, perceptions of risk, and changing attitudes towards HIV/AIDS: New evidence from a longitudinal study using fixed-effects analysis. *Popul. Stud. (NY).*, 59(3):265–282.

[100] T. Herath and H. Rao. Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness. *Decis. Support Syst.*, 47(2):154–165, May 2009.

[101] C. Herley and D. Florêncio. Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. In *Econ. Inf. Secur. Priv.*, pages 33–53. Springer, 2010.

[102] C. Herley, P. van Oorschot, and A. S. Patrick. Passwords: If We're So Smart, Why Are We Still Using Them? In R. Dingledine and P. Golle, editors, *Financ. Cryptogr. Data Secur. – 13th Int. Conf. 2009*, volume 5628 of *LNCS*, pages 230–237, 2009.

[103] H. W. Hethcote, J. W. Van Ark, and I. M. Longini. A simulation model of AIDS in San Francisco: I. model formulation and parameter estimation. *Math. Biosci.*, 106(2):203–222, Oct. 1991.

[104] P. F. Horwood, S. Karl, I. Mueller, M. H. Jonduo, B. I. Pavlin, R. Dagina, B. Ropa, S. Bieb, A. Rosewell, M. Umezaki, P. M. Siba, and A. R. Greenhill. Spatio-temporal epidemiology of the cholera outbreak in Papua New Guinea, 2009-2011. *BMC Infect. Dis.*, 14:449, Jan. 2014.

[105] N. Husted and S. Myers. Why Mobile-to-Mobile Wireless Malware Won't Cause a Storm. In *Proc. 4th USENIX Conf. Large-scale Exploit. emergent Threat.* USENIX Association, 2011.

[106] T. N. Jagatic, N. A. Johnson, M. Jakobsson, and F. Menczer. Social phishing. *Commun. ACM*, 50(10):94–100, Oct. 2007.

[107] T. Jim, N. Swamy, and M. Hicks. Defeating script injection attacks with browser-enforced embedded policies. In *Proc. 16th Int. Conf. World Wide Web - WWW '07*, page 601, New York, New York, USA, 2007. ACM Press.

[108] A. C. Johnston and M. Warkentin. FEAR APPEALS AND INFORMATION SECURITY BEHAVIORS: AN EMPIRICAL STUDY. *MIS Q.*, 34(3):549–566, 2010.

[109] J. Johnston, J. Eloff, and L. Labuschagne. Security and human computer interfaces. *Comput. Secur.*, 22(8):675–684, 2003.

[110] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxon, and S. Savage. Spamalytics: An empirical analysis of spam marketing conversion. In *Proc. 15th ACM Conf. Comput. Commun. Secur.*, pages 3–14, Alexandria, VA, USA, 2008. ACM.

[111] M. Katz and C. Shapiro. Systems competition and network effects. *J. Econ. Perspect.*, 8(2):93—-115, 1994.

[112] T. Kelley and L. J. Camp. Online promiscuity: Prophylactic patching and the spread of computer transmitted infections. In *Work. Econ. Inf. Secur. (WEIS' 12)*, Berlin, Germany, 2012.

[113] J. Kephart and S. White. Directed-graph epidemiological models of computer viruses. *Proceedings. 1991 IEEE Comput. Soc. Symp. Res. Secur. Priv.*, 0:343–359, 1991.

[114] J. Kephart and S. White. Measuring and modeling computer virus prevalence. *Proc. 1993 IEEE Comput. Soc. Symp. Res. Secur. Priv.*, 0:2–15, 1993.

[115] W. O. Kermack and A. G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. *Proc. R. Soc. London. Ser. A, Contain. Pap. a Math. Phys. Character*, 115(772):700–721, Aug. 1927.

[116] I. Kirlappos, M. Sasse, and N. Harvey. Why trust seals don't work: A study of user perceptions and behavior. In S. Katzenbeisser, E. Weippl, L. J. Camp, M. Volkamer, M. Reiter, and X. Zhang, editors, *Trust Trust. Comput.*, pages 308–324. Springer Berlin Heidelberg, 2012.

[117] S. Kitchovitch and P. Liò. Risk perception and disease spread on social networks. *Procedia Comput. Sci.*, 1(1):2345–2354, May 2010.

[118] J. Knight, M. Elder, and C. Wang. On computer viral infection and the effect of immunization. In *Proc. 16th Annu. Comput. Secur. Appl. Conf.*, pages 246–256. IEEE Comput. Soc, 2000.

[119] M. Konte, N. Feamster, and J. Jung. Dynamics of online scam hosting infrastructure. *Passiv. Act. Netw. Meas.*, (1):219–228, 2009.

[120] S. Kraemer, P. Carayon, and J. Clem. Human and organizational factors in computer and information security: Pathways to vulnerabilities. *Comput. Secur.*, 28(7):509–520, Oct. 2009.

[121] M. KRASSA. Social groups, selective perception, and behavioral contagion in public opinion. *Soc. Networks*, 10(2):109–136, June 1988.

[122] B. Krebs. Google: Your Computer Appears to Be Infected, 2011.

[123] B. Krebs. Scammers Swap Google Images for Malware, 2011.

[124] B. Krebs. Spam Volumes: Past & Present, Global & Local, 2013.

[125] P. Kusev, P. van Schaik, P. Ayton, J. Dent, and N. Chater. Exaggerated risk: prospect theory and probability weighting in risky choice. *J. Exp. Psychol. Learn. Mem. Cogn.*, 35(6):1487–505, Nov. 2009.

[126] A. Lahiri. Revisiting the Incentive to Tolerate Illegal Distribution of Software Products. In *44th Hawaii Int. Conf. Syst. Sci.*, 2011.

[127] W. Liambila, I. Askew, J. Mwangi, R. Ayisi, J. Kibaru, and S. Mullick. Feasibility and effectiveness of integrating provider-initiated testing and counselling within family planning services in Kenya. *AIDS*, 23 Suppl 1:S115–21, Nov. 2009.

[128] J. Lin, N. Sadeh, S. Amini, J. Lindqvist, J. I. Hong, and J. Zhang. Expectation and purpose. In *Proc. 2012 ACM Conf. Ubiquitous Comput. - UbiComp '12*, page 501, New York, New York, USA, 2012. ACM Press.

[129] D. Liu, X. Wang, and L. J. Camp. Mitigating Inadvertent Insider Threats with Incentives. In R. Dingledine and P. Golle, editors, *Financ. Cryptogr. Data Secur.*, pages 1–16. Springer Berlin Heidelberg, 2009.

[130] J. Lorince, D. Donato, and P. Todd. Path Following in Social Web Search. In W. Kennedy, N. Agarwal, and S. Yang, editors, *Soc. Comput. Behav. Model. Predict. SE - 15*, volume 8393 of *Lecture Notes in Computer Science*, pages 119–127. Springer International Publishing, 2014.

[131] E. Maiberg. Rovio brings Angry Birds Friends to mobile, 2013.

[132] T. Maillart and D. Sornette. Heavy-tailed distribution of cyber-risks. *Eur. Phys. J. B*, 75(3):357–364, Apr. 2010.

[133] D. Malone and K. Maher. Investigating the distribution of password choices. *Proc. 21st Int. Conf. World Wide Web - WWW '12*, page 301, 2012.

[134] J. N. Marewski, W. Gaissmaier, and G. Gigerenzer. Good judgments do not require complex cognition. *Cogn. Process.*, 11(2):103–21, May 2010.

[135] S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.*, 254(1):178–96, Sept. 2008.

[136] G. McDonald and B. Hope. Deceptive Downloads: "Clean Downloads" Turn Nasty, 2014.

[137] M. R. Meiss, F. Menczer, S. Fortunato, A. Flammini, and A. Vespignani. Ranking web sites with real user traffic. *Proc. Int. Conf. Web search web data Min. - WSDM '08*, page 65, 2008.

[138] M. R. Meiss, F. Menczer, and A. Vespignani. Structural analysis of behavioral networks from the Internet. *J. Phys. A Math. Theor.*, 41(22):224022, June 2008.

[139] J. Mickens and B. Noble. Modeling epidemic spreading in mobile environments. In *Proc. 4th ACM Work. Wirel. Secur.* ACM, 2005.

[140] G. Miritello, E. Moro, R. Lara, R. Martínez-López, J. Belchamber, S. G. Roberts, and R. I. Dunbar. Time as a limited resource: Communication strategy in mobile phone networks. *Soc. Networks*, 35(1):89–95, Jan. 2013.

[141] C. D. Moore, M. X. Cohen, and C. Ranganath. Neural mechanisms of expert skills in visual working memory. *J. Neurosci.*, 26(43):11187–96, Oct. 2006.

[142] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver. Inside the slammer worm. *IEEE Secur. Priv. Mag.*, 1(4):33–39, July 2003.

[143] D. Moore, C. Shannon, and J. Brown. Code-Red: a case study on the spread and victims of an Internet worm. In *Proc. 2nd ACM SIGCOMM Work. Internet Meas.*, pages 273–284. ACM, 2002.

[144] D. Moore, C. Shannon, G. Voelker, and S. Savage. Internet quarantine: requirements for containing self-propagating code. In *IEEE INFOCOM 2003. Twenty-second Annu. Jt. Conf. IEEE Comput. Commun. Soc. (IEEE Cat. No.03CH37428)*, volume 3, pages 1901–1910. IEEE, 2003.

[145] T. Moore and R. Clayton. Evil searching: Compromise and recompromise of internet hosts for phishing. *Financ. Cryptogr. Data Secur.*, pages 256–272, 2009.

[146] T. Moore, R. Clayton, and R. Anderson. The Economics of Online Crime. *J. Econ. Perspect.*, 23(3):3–20, Aug. 2009.

[147] T. Moore, R. Clayton, and H. Stern. Temporal correlations between spam and phishing websites. *2nd USENIX Work. Large-Scale*, 2009.

[148] M. MORRIS. Epidemiology and Social Networks: Modeling Structured Diffusion. *Sociol. Methods Res.*, 22(1):99–126, Aug. 1993.

[149] M. Moslonka-Lefebvre, S. Bonhoeffer, and S. Alizon. Weighting for sex acts to understand the spread of STI on networks. *J. Theor. Biol.*, 311:46–53, Oct. 2012.

[150] M. Moussaïd. Opinion formation and the collective dynamics of risk perception. *PLoS One*, 8(12):e84592, Jan. 2013.

[151] M. Moussaïd, J. E. Kämmer, P. P. Analytis, and H. Neth. Social influence and the collective dynamics of opinion formation. *PLoS One*, 8(11):e78433, Jan. 2013.

[152] P. Mozur. Protesters in Hong Kong are Targets of Scrutiny Through Their Phones, Oct. 2014.

[153] S. Myers. Introduction to Phishing. In M. Jakobsson and S. Myers, editors, *Phishing Countermeas.*, pages 1–29. John Wiley & Sons, Inc., 2006.

[154] S. Nagaraja, P. Mittal, C. Hong, M. Caesar, and N. Borisov. BotGrep: Finding Bots with Structured Graph Analysis. *USENIX 2010*, pages 1–24, 2010.

[155] A. Nematzadeh, E. Ferrara, A. Flammini, and Y.-y. Ahn. Optimal network clustering for information diffusion. page 5, Jan. 2014.

[156] H. Neth, B. Meder, A. Kothiyal, and G. Gigerenzer. Homo heuristicus in the financial world: From risk management to managing uncertainty. *J. Risk Manag. Financ. Institutions*, 7(2):134—-144., 2014.

[157] M. Newman, S. Forrest, and J. Balthrop. Email networks and the spread of computer viruses. *Phys. Rev. E*, 66(3):035101, Sept. 2002.

[158] G. Norcie, J. Blythe, K. Caine, and L. J. Camp. Why Johnny Can't Blow the Whistle: Identifying and Reducing Usability Issues in Anonymity Systems. *Proc. 2014 Work. Usable Secur.*, 2014.

[159] OpenID Foundation. Specifications, 2010.

[160] S. Oviatt. Human-centered design meets cognitive load theory. In *Proc. 14th Annu. ACM Int. Conf. Multimed. - Multimed. '06*, page 871, New York, New York, USA, 2006. ACM Press.

[161] W. Pan, N. Aharony, and A. Pentland. Composite social network for predicting mobile apps installation. In *AAAI*, 2011.

[162] A. Pashalidis and C. J. Mitchell. A Taxonomy of Single Sign-On Systems. In R. Safavi-Naini and J. Seberry, editors, *Proc. 8th Australas. Conf. Inf. Secur. Priv. 2003*, volume 2727 of *LNCS*. Springer, 2003.

[163] R. Pastor-Satorras and A. Vespignani. Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E*, 63(6):1–8, May 2001.

[164] R. Pastor-Satorras and A. Vespignani. Epidemic Spreading in Scale-Free Networks. *Phys. Rev. Lett.*, 86(14):3200–3203, Apr. 2001.

[165] A. Pathak, F. Qian, C. Y. Hu, M. Z. Mao, and R. Supranamaya. Botnet spam campaigns can be long lasting: evidence, implications, and analysis. In *Proc. Elev. Int. Jt. Conf. Meas. Model. Comput. Syst.*, pages 13–24, Seattle, WA, USA, 2009. ACM.

[166] A. Patrick. Commentary on research on new security indicators, March 2007.

[167] A. S. Patrick, A. C. Long, and S. Flinn. HCI and security systems. In *CHI '03 Ext. Abstr. Hum. factors Comput. Syst. - CHI '03*, page 1056, New York, New York, USA, 2003. ACM Press.

[168] N. Perra, D. Balcan, B. Gonçalves, and A. Vespignani. Towards a characterization of behavior-disease models. *PLoS One*, 6(8):e23084, Jan. 2011.

[169] H. Pilz and S. Schindler. Are free Android virus scanners any good. Technical Report November, AVTest, Magdeburg Germany, 2011.

[170] P. S. Press. Cyber-crooks manipulate Internet searches to sell fake antivirus products, 2009.

[171] F. Raja, K. Hawkey, and K. Beznosov. Revealing hidden context. In *Proc. 5th Symp. Usable Priv. Secur. - SOUPS '09*, page 1, New York, New York, USA, 2009. ACM Press.

[172] A. Rao, B. Jha, and G. Kini. Effect of grammar on security of long passwords. In *Proc. third ACM Conf. Data Appl. Secur. Priv. - CODASPY '13*, page 317, New York, New York, USA, 2013. ACM Press.

[173] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, A. Flammini, and F. Menczer. Detecting and Tracking Political Abuse in Social Media. In *Proc. 5th Int. AAAI Conf. Weblogs Soc. Media*, 2011.

[174] J. Ratkiewicz, A. Flammini, and F. Menczer. Traffic in Social Media I: Paths Through Information Networks. In *2010 IEEE Second Int. Conf. Soc. Comput.*, pages 452–458. IEEE, Aug. 2010.

[175] J. Ratkiewicz, S. Fortunato, A. Flammini, F. Menczer, and A. Vespignani. Characterizing and Modeling the Dynamics of Online Popularity. *Phys. Rev. Lett.*, 105(15):158701, Oct. 2010.

[176] a. M. Renton, L. Whitaker, and M. Riddlesdell. Heterosexual HIV transmission and STD prevalence: predictions of a theoretical model. *Sex. Transm. Infect.*, 74(5):339–344, Oct. 1998.

[177] H.-S. Rhee, C. Kim, and Y. U. Ryu. Self-efficacy in information security: Its influence on end users' information security practice behavior. *Comput. Secur.*, 28(8):816–826, Nov. 2009.

[178] RSA FraudAction Research Labs. Anatomy of an Attack.

[179] A. Saita. Mobile Malware Dubbed 'Bill Shocker' Targets Chinese Android Users, 2013.

[180] M. Sasse, S. Brostoff, and D. Weirich. Transforming the 'weakest link'—a human/computer interaction approach to usable and effective security. *BT Technol. J.*, 19(3):122–131, 2001.

[181] S. Schechter, R. Dhamija, A. Ozment, and I. Fischer. The Emperor's New Security Indicators. *2007 IEEE Symp. Secur. Priv. (SP '07)*, 0:51–65, May 2007.

[182] S. Schechter, R. Dhamija, A. Ozment, and I. Fischer. The Emperor's New Security Indicators. *2007 IEEE Symp. Secur. Priv. (SP '07)*, 0:51–65, May 2007.

[183] S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer. The emperor's new security indicators: An evaluation of website authentication and the effect of role playing on usability studies. In *Proc. IEEE Symposium on Security and Privacy (S&P) 2007*, pages 51–65. IEEE Press, 2007.

[184] Science and Technology Committee. Personal Internet Security. In *5th Rep. Sess. 2006-2007*, volume I of *5th Repo*, page 121. House of Lords, House of Lords, 2007.

[185] G. Serazzi and S. Zanero. Computer Virus Propagation Models. In *Perform. Tools Appl. to Networked Syst.*, pages 26–50. Springer, Berlin, Germany, 2004.

[186] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs. Who falls for phish? In *Proc. 28th Int. Conf. Hum. factors Comput. Syst. - CHI '10*, page 373, New York, New York, USA, 2010. ACM Press.

[187] S. Shiboski and N. S. Padian. Population- And Individual-Based Approaches To The Design And Analysis Of Epidemiologic Studies Of Sexually Transmitted Disease Transmission. *J. Infect. Dis.*, 174(Supplement 2):S188–S200, Oct. 1996.

[188] H. A. Simon. Rational choice and the structure of the environment. *Psychol. Rev.*, 63(2):129–38, Mar. 1956.

[189] H. A. Simon. Invariants of human behavior. *Annu. Rev. Psychol.*, 41:1–19, 1990.

[190] A. Smith. Nearly half of American adults are smartphone owners, 2012.

[191] J. Sobey, R. Biddle, P. van Oorschot, and A. S. Patrick. Exploring User Reactions to New Browser Cues for Extended Validation Certificates. In S. Jajodia and J. Lopez, editors, *Proc. 13th Eur. Symp. Res. Comput. Secur. 2008*, volume 5283 of *LNCS*, pages 411–427. Springer, 2008.

[192] SOMEDROID. F-Secure report shows once again why you should stick to the Play Store for app downloads, 2014.

[193] A. Sotirakopoulos, K. Hawkey, and K. Beznosov. "I did it because I trusted you": Challenges with the study environment biasing participant behaviours. In *SOUPS Usable Security Experiment Reports (USER) Workshop*, 2010.

[194] A. Sotirakopoulos, K. Hawkey, and K. Beznosov. On the challenges in usable security lab studies: Lessons learned from replicating a study on SSL warnings. In L. F. Cranor, editor, *Proc. 7th Symposium on Usable Privacy and Security (SOUPS) 2011*. ACM, 2011.

[195] J. E. Spain, J. F. Peipert, T. Madden, J. E. Allsworth, and G. M. Secura. The Contraceptive CHOICE Project: recruiting women at highest risk for unintended pregnancy and sexually transmitted infection. *J. Women's Heal.*, 19(12):2233–8, Dec. 2010.

[196] S. Staniford, V. Paxson, and N. Weaver. How to own the internet in your spare time. In *Proc. 11th USENIX Secur. Symp.*, volume 8, pages 149–167, 2002.

[197] J. M. Stanton, K. R. Stam, P. Mastrangelo, and J. Jolton. Analysis of end user security behaviors. *Comput. Secur.*, 24(2):124–133, Mar. 2005.

[198] D. Stebila. Reinforcing bad behaviour. In *Proc. 22nd Conf. Comput. Interact. Spec. Interes. Gr. Aust. Comput. Interact. - OZCHI '10*, page 248, New York, New York, USA, 2010. ACM Press.

[199] D. Stebila. Reinforcing bad behaviour: the misuse of security indicators on popular websites. In *Proc. 22nd Australasian Conf. on Computer-Human Interaction (OzCHI) 2010*, pages 248–251. ACM, 2010.

[200] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna. Your botnet is my botnet: Analysis of a botnet takeover. In *Proc. 16th ACM Conf. Comput. Commun. Secur.*, pages 635–647. ACM, 2009.

[201] S.-T. Sun, E. Pospisil, I. Muslukhov, N. Dindar, K. Hawkey, and K. Beznosov. What makes users refuse web single sign-on?: an empirical investigation of {OpenID}. In L. F. Cranor, editor, *Proc. 7th Symp. Usable Priv. Secur. 2011*, pages 4:1—-4:20. ACM, 2011.

[202] J. Sunshine, S. Egelman, H. Almuhimedi, N. Atri, and L. F. Cranor. Crying Wolf: An Empirical Study of {SSL} Warning Effectiveness. In *Proc. 18th {USENIX} Secur. Symp.*, 2009.

[203] C. Szongott, B. Henne, and M. Smith. Evaluating the threat of epidemic mobile malware. In *Wirel. Mob. Comput. Netw. Commun. (WiMob), 2012 IEEE 8th Int. Conf.*, pages 443–450. IEEE, 2012.

[204] K. Thomas and D. M. Nicol. The Koobface botnet and the rise of social malware. In *2010 5th Int. Conf. Malicious Unwanted Softw.*, pages 63–70. IEEE, Oct. 2010.

[205] P. Thomas, J. Ryan, and T. Cook. Guilty Plea in Los Alamos Security Breach, 2007.

[206] A. Tsow and L. Camp. A privacy-aware architecture for sharing web histories. *IBM Syst. J.*, (0705676):1–13, 2007.

[207] N. B. Turk-Browne, J. Jungé, and B. J. Scholl. The automaticity of visual statistical learning. *J. Exp. Psychol. Gen.*, 134(4):552–64, Nov. 2005.

[208] N. B. Turk-Browne, B. J. Scholl, M. M. Chun, and M. K. Johnson. Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *J. Cogn. Neurosci.*, 21(10):1934–45, Oct. 2009.

[209] N. B. Turk-Browne, B. J. Scholl, M. K. Johnson, and M. M. Chun. Implicit perceptual anticipation triggered by statistical learning. *J. Neurosci.*, 30(33):11177–87, Aug. 2010.

[210] R. Turn. Security and Privacy Requirements in Computing. In *Proc. 1986 ACM Fall Jt. Comput. Conf.*, volume 54, pages 1106–1114. IEEE Computer Society Press, 1986.

[211] M. Van Eeten, J. Bauer, J. Groenewegen, W. Lemstra, and M. van Eeten. The economics of malware. In *Proc. 35th Telecommun. Policy Res. Conf.*, volume 31, Arlington, CA, 2007.

[212] P. van Schaik and J. Ling. A cognitive-experiential approach to modelling web navigation. *Int. J. Hum. Comput. Stud.*, 70(9):630–651, Sept. 2012.

[213] K. E. Vaniea, E. Rader, and R. Wash. Betrayed by updates. In *Proc. 32nd Annu. ACM Conf. Hum. factors Comput. Syst. - CHI '14*, pages 2671–2674, New York, New York, USA, 2014. ACM Press.

[214] H. Varian, F. Wallenberg, and G. Woroch. Who signed up for the do-not-call list? *Communications*, 89:1–26, 2004.

[215] C. Viecco, A. Tsow, and L. J. Camp. A privacy-aware architecture for a Web rating system. *IBM J. Res. Dev.*, 53(2):7:1–7:16, Mar. 2009.

[216] T. Vila, R. Greenstadt, and D. Molnar. Why we can't be bothered to read privacy policies models of privacy economics as a lemons market. In *Proc. 5th Int. Conf. Electron. Commer. - ICEC '03*, pages 403–407, New York, New York, USA, 2003. ACM Press.

[217] P. Wang, M. C. González, R. Menezes, and A.-L. Barabási. Understanding the spread of malicious mobile-phone programs and their damage potential. *Int. J. Inf. Secur.*, 12(5):383–392, June 2013.

[218] Y. Wang and C. Wang. Modeling the effects of timing parameters on virus propagation. In *Proc. 2003 ACM Work. Rapid Malcode - WORM'03*, page 61, New York, New York, USA, 2003. ACM Press.

[219] R. Wash. Folk Models of Home Computer Security. In *Symp. Usable Priv. Secur.* ACM Press, 2010.

[220] R. Wash, E. Rader, K. Vaniea, and M. Rizor. Out of the Loop: How Automated Software Updates Cause Unintended Security Consequences. In *Proc. 10th Symp. Usable Priv. Secur.*, Menlo Park, CA, 2014. USENIX Association.

[221] N. Weaver, V. Paxson, S. Staniford, and R. Cunningham. A Taxonomy of Computer Worms. In *Proc. 2003 ACM Work. Rapid Malcode*, pages 11–18. ACM, 2003.

[222] M. Weir, S. Aggarwal, M. Collins, and H. Stern. Testing metrics for password creation policies by attacking large sets of revealed passwords. *Proc. 17th ACM Conf. Comput. Commun. Secur. - CCS '10*, page 162, 2010.

[223] L. Weng, F. Menczer, and Y.-Y. Ahn. Virality prediction and community structure in social networks. *Sci. Rep.*, 3:2522, Jan. 2013.

[224] L. Weng, J. Ratkiewicz, N. Perra, B. Gonçalves, C. Castillo, F. Bonchi, R. Schifanella, F. Menczer, and A. Flammini. The role of information diffusion in the evolution of social networks. In *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '13*, page 356, New York, New York, USA, 2013. ACM Press.

[225] T. Whalen and K. M. Inkpen. Gathering evidence: use of visual security cues in web browsers. In K. M. Inkpen and M. van de Panne, editors, *Proc. Graphics Interface 2005*, volume 112 of *Graphics Interface*, pages 137–144. Canadian Human-Computer Communications Society, 2005.

[226] A. Whitten and J. D. Tygar. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0. In *USENIX Secur.*, 1999.

[227] T. N. Wisdom and R. L. Goldstone. Innovation, imitation, and problem-solving in a networked group. *Nonlinear Dynamics. Psychol. Life Sci.*, 15(2):229–52, Apr. 2011.

[228] T. N. Wisdom, X. Song, and R. L. Goldstone. Social learning strategies in networked groups. *Cogn. Sci.*, 37(8):1383–425, 2013.

[229] M. Workman. Gaining Access with Social Engineering: An Empirical Study of the Threat. *Inf. Syst. Secur.*, 16(6):315–331, Dec. 2007.

[230] M. Workman, W. H. Bommer, and D. Straub. Security lapses and the omission of information security measures: A threat control model and empirical test. *Comput. Human Behav.*, 24(6):2799–2816, Sept. 2008.

[231] T. Worstall. There Is Too Malware On The iPhone!, 2012.

[232] M. J. Wright, D. T. Bishop, R. C. Jackson, and B. Abernethy. Functional MRI reveals expert-novice differences during sport-related anticipation. *Neuroreport*, 21(2):94–98, Jan. 2010.

[233] R. T. Wright and K. Marett. The influence of experiential and dispositional factors in phishing: An empirical investigation of the deceived. *J. Management Info. Sys.*, 27(1):273–303, July 2010.

[234] M. Wu, R. C. Miller, and S. L. Garfinkel. Do security toolbars actually prevent phishing attacks? In *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI '06*, page 601, New York, New York, USA, 2006. ACM Press.

[235] C. Xiao. APPBUYER: NEW IOS MALWARE STEALS APPLE ID AND PASSWORD TO BUY APPS, 2014.

[236] S.-H. Yook, H. Jeong, and A.-L. Barabasi. Modeling the Internet's large-scale topology. *Proc. Natl. Acad. Sci. U. S. A.*, 99(21):13382–6, Oct. 2002.

[237] J. Zhao, N. Al-Aidroos, and N. B. Turk-Browne. Attention is spontaneously biased toward regularities. *Psychol. Sci.*, 24(5):667–77, May 2013.

[238] Y. Zhou, Z. Wang, W. Zhou, and X. Jiang. Hey, you, get off of my market: Detecting malicious apps in official and alternative Android markets. In *Proc. 19th Annu. Netw. Distrib. Syst. Secur. Symp.*, number 2, 2012.

[239] C. Zou and D. Towsley. Routing Worm: A Fast, Selective Attack Worm Based on IP Address Information. In *Work. Princ. Adv. Distrib. Simul.*, pages 199–206. IEEE, 2005.

[240] C. C. Zou, W. Gong, and D. Towsley. Code red worm propagation modeling and analysis. In *Proc. 9th ACM Conf. Comput. Commun. Secur. - CCS '02*, page 138, New York, New York, USA, 2002. ACM Press.

[241] C. C. Zou, W. Gong, and D. Towsley. Worm propagation modeling and analysis under dynamic quarantine defense. In *Proc. 2003 ACM Work. Rapid Malcode - WORM'03*, page 51, New York, New York, USA, 2003. ACM Press.

Curriculum Vitae

Education

Indiana University, Bloomington, IN

Ph.D. Informatics and Cognitive Science (2014)
Thesis: *Systemic Effects of Human Factors in Information Security*
Advisors: L. Jean Camp, Alessandro Flammini, Rob Goldstone, and Peter M. Todd

M.Sc. Bioinformatics (2009)
Thesis: *Periodicity of Notch oscillating signaling pathway is controlled by the topology of its regulatory networks during somite formation*
Project: Discrete Network Dynamics and Extract and Annotate Attractors for Network Bench
Advisors: Mehmet (Memo) Dalkilc and Santiago Schnell

James Madison University, Harrisonburg, VA

M.Sc. Information Security (2007)
Thesis: *A Method for Increasing Transmission Rates in Covert Timing Channels*
Advisor: Steven J. Greenwald

Calvin College, Grand Rapids, MI

B.Sc. Computer Science (2001)

Research Experience

Research Assistant L. Jean Camp, Indiana University, Bloomington, IN (2013 - 2014)

- Developed and implemented an agent-based model of smart phone users' application adoption strategies and marketplace configurations to study their effects on malicious mobile application spread.
- Analyzed the model results using a Bayesian hierarchical model to examine the overall importance of the different factors in mitigating malicious mobile application spread.
- Developed an algorithm for analyzing cascading attack and fault tolerance in multi-level networks for comparing the robustness of software defined networks and BGP-type networks.

Visiting Research Assistant Andrew Adams and Kiyoshi Murata, Meiji University, Tokyo, Japan (2013)

- Worked with an international team to adapt my eye-tracking experimental design to Japanese subjects.
- Ran the adjusted experimental design to evaluate the external validity of previous eye-tracking results.

- Participated as a member of the program committee for The Workshop on Usable Security 2013, Okinawa, Japan.

REU Undergraduate Research Mentor L. Jean Camp, Indiana University, Bloomington, IN (2011)

- Supervised undergraduate research fellow in experimental design
- Executed an eye-tracking experiment on graduate students at the School of Informatics and Computing to evaluate attention to web browser-based security cues in self-identified expert users.
- Initial results accepted at LASER 2012.

Research Assistant Jim Sherman and Alessandro Vespignani, Indiana University, Bloomington, IN (2008-2011)

- Developed an epidemiological model of adolescent smoking adoption.
- Collected available social contagion data sets for inclusion in the EPiC–A cyberinfrastructure for epidemiological research and data sharing–database.

Software Developer Network Workbench, Indiana University, Bloomington, IN (2007-2008)

- Developed data converters to allow for multiple network description formats to be understood by Network Workbench.
- Designed and implemented a suite of tools–Discrete Network Dynamics and Extract and Annotate Attractors–for the analysis of discrete network model dynamics in Network Workbench.
- Implemented algorithms for the extraction of several types of networks from common scholarly database formats.

Research Assistant Institute for Infrastructure and Information Assurance, Harrisonburg, VA (2004-2006)

- Researched implementations of covert channels in UDP.
- Designed and implemented a software-based UDP timing channel in TFTP, and, using a method for bit packing, created a more effective covert channel than the standard symmetric timing channel.

REU Research Assistant Hope College, Holland, MI (2001)

- Investigated possible cryptographic key infrastructures for digital rights management in e-textbooks.
- Designed a protoype system using Java, `XrML`, and password-pased cryptography to protect both fair-use and intellectual property concerns.

Publications/Presentations/Posters

Kelley, T. Ecology of Security: Brains, Bodies, and Environments. *22nd USENIX Security Symposium*, (2013).

Arianezhad, M., Camp, L. J., Kelley, T., & Stebila, D. Comparative eye tracking of experts and novices in web single sign-on. *Proceedings of the third ACM Conference on Data and Application Security and Privacy*, (2013).

Kelley, T., Camp, L. J., Lien, S. & Stebila, D. Self-identified experts lost on the interwebs. *Proceedings of the 2012 Workshop on Learning from Authoritative Security Experiment Results*, (2012).

Kelley, T. & Camp, L. J. Online promiscuity: Prophylactic patching and the spread of computer transmitted infections. *Workshop on the Economics of Information Security*, (2012).

Kelleyt, T. Media Context and Source Credibility: A Pilot Study. *Future Internet and Society: A Complex Systems Perspective*, (2010).

Bubenheim, D. & Kelley, T. Data Integration and Visualization. *iPlant Grand Challenge Workshop: Mechanistic Basis of Plant Adaption*, (2008).

Kelley, T. Discrete Network Dynamics Analysis Using Network Workbench. *Systems Biology Dynamics: from Genes to Organisms*, (2008).

Kelley, T. A Method for Increasing Transmission Rates in Covert Timing Channels. *James Madison University, Department of Computer Science*, (2008).

Huang, W., Börner, K., Duhon, R. J., Kelley, T., Herr II, B. W. & Schnell, S. NWB workshop: Towards an All-in-One Tool for Network Scientists Interested in Large Scale Network Analysis, Modeling, and Visualization. *Talk Series of Network and Complex Systems*, (2007).

Teaching Experience

Lecturer I230 – Analytical Foundations of Security, Indiana University, Bloomington (Fall 2013)

Lecturer I433/533 CS629 – Systems, Protocol Security & Information Assurance, Indiana University, Bloomington (Fall 2013)

Associate Instructor I330 – Legal and Organization Security Informatics, Indiana University, Bloomington (Spring 2012)

Associate Instructor I202 – Social Informatics, Indiana University, Bloomington (Fall 2011)

Lecturer Tips for Peer Teaching, Indiana University, Bloomington (Summer 2010)

Lecturer Beyond Plug and Chug: Active Learning and Critical Thinking in Life, Physical, and Social Sciences (Summer 2010)

Associate Instructor I201 – Mathematical Foundations of Informatics, Indiana University, Bloomington (2006-2007)

Professional Development

McGill Summer Sessions in Systems Biology, McGill University, Montreal (2008)

NSA Certified Information Systems Security Professional–NSTISSI No. 4011 (2007)

NSA Certified Information Systems Security Officer–CNSSI No. 4014 (2007)

Awards and Honors

LASER Workshop Student Travel Grant Award (2013)

Volkswagen Foundation Travel Grant Award (2012)

REU Undergraduate Research Fellow Mentor (2011)

The Dynamics of Brain-Body-Environment Systems in Behavior and Cognition IGERT Associate

The Dynamics of Brain-Body-Environment Systems in Behavior and Cognition IGERT Summer Fellowship Grant (2011)

ESF-COST Travel Grant

REU Undergraduate Research Fellow